



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 121196

TO: Manjunath N Rao
Location: rem/3b11/3c70
Art Unit: 1652
Friday, May 14, 2004
Case Serial Number: 10/016725

From: Paul Schulwitz
Location: Biotech-Chem Library
REM-1A65
Phone: (571)272-2527

paul.schulwitz@uspto.gov

Search Notes

Examiner Rao,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
STIC Biotech/Chem Library
(571)272-2527

STIC-Biotech/ChemLib

121196

From: Rao, Manjunath N.
Sent: Wednesday, May 05, 2004 10:22 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 10/016,725

From: Manjunath N. Rao
Art Unit 1652, Room ~~10711~~
Mail Box in Room ~~10711~~
Phone: 306-5681

Date: 5-5-04

Please search the following as soon as possible for application with serial number **10/016725**

1. Nucleotides 1-3958 of SEQ ID NO: 15

2. Nucleotides 1-3910 of SEQ ID NO:15

against all commercial nucleic acid databases including issued patents database and pending application database and provide a print of all results.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao
Art Unit 1652, Room 3B81,
Mail Box in Room 3C70,
Remsen Building, USPTO
400, Dulany St.
Alexandria, VA.
Phone: 571-272-0939

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 5/4
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:19 ; Search time 23820.8 Seconds
(without alignments)
7114.428 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3910
Perfect score: 3910
Sequence: 1 aggaattcattcatttaataa.....cggacacggtcaccogctgc 3910

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pla.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3910	100.0	4150	6	AX528729	Sequence
2	3910	100.0	4150	9	AF320307	AF320307 Homo sapi
C 3	3880.2	99.2	159351	2	AC018783	AC018783 Homo sapi
C 4	3869.8	99.0	115278	9	AL139819	AL139819 Human DNA
C 5	1779.2	45.5	185644	2	AL391723	AL391723 Homo sapi
C 6	1751.8	44.8	1970	6	AX431388	AX431388 Sequence
7	1711.6	43.8	161827	2	AC139668	AC139668 Papio ham
8	1448.2	37.0	179879	2	AC026883	AC026883 Homo sapi
9	778.2	19.9	190784	2	AC139669	AC139669 Lemur cat
10	500	12.8	17088	4	AX241932	AY241932 Bos tauru
11	460	11.8	164201	9	AL157935	AL157935 Human DNA
12	450.4	11.5	617	6	AX357255	AX357255 Sequence
13	447.2	11.4	200288	2	AC119725	AC119725 Homo sapi
14	433	11.1	175842	9	AL731547	AL731547 Human DNA
C 15	423.8	10.8	190508	9	AC005037	AC005037 Homo sapi
C 16	415.2	10.6	44496	9	AC004760	AC004760 Homo sapi
17	415.2	10.6	190814	9	AC006111	AC006111 Homo sapi
C 18	408.8	10.5	183690	9	AC091180	AC091180 Homo sapi
19	407.4	10.4	170528	9	AC083866	AC083866 Homo sapi
20	407.4	10.4	217615	9	AC011005	AC011005 Homo sapi
C 21	404.2	10.3	137289	9	AL161909	AL161909 Human DNA
C 22	403	10.3	189179	2	AC090265	AC090265 Homo sapi
C 23	402.8	10.3	181597	9	AC090515	AC090515 Homo sapi
C 24	402.8	10.3	187734	9	AC025918	AC025918 Homo sapi
C 25	401.8	10.3	204963	2	AC145617	AC145617 Homo sapi
C 26	394.2	10.1	179599	2	AC013278	AC013278 Homo sapi
27	394.2	10.1	195558	9	AC007448	AC007448 Homo sapi
C 28	392.2	10.0	155559	2	AL133336	AL133336 Homo sapi
C 29	392.2	10.0	169835	9	AP005717	AP005717 Homo sapi
C 30	391.6	10.0	171641	9	AC018512	AC018512 Homo sapi
C 31	390	10.0	139632	9	AC145334	AC145334 Pan trogl
C 32	390	10.0	171555	2	AC087537	AC087537 Homo sapi
C 33	390	10.0	181891	9	AC023356	AC023356 Homo sapi
C 34	389.8	10.0	67199	9	AC025162	AC025162 Homo sapi
C 35	388.4	9.9	172421	2	AC013279	AC013279 Homo sapi
C 36	386.2	9.9	112364	9	AC109927	AC109927 Homo sapi
C 37	386	9.9	73162	9	AC105095	AC105095 Homo sapi
C 38	386	9.9	197092	2	AC036155	AC036155 Homo sapi
C 39	384.8	9.8	140529	2	AC011486	AC011486 Homo sapi
40	384	9.8	133475	9	AC006254	AC006254 Homo sapi
41	382.4	9.8	189639	9	AC010997	AC010997 Homo sapi
C 42	381.6	9.8	80775	9	AC068539	AC068539 Homo sapi
C 43	381.2	9.7	150681	9	AC006011	AC006011 Homo sapi
44	381.2	9.7	187064	9	AC011476	AC011476 Homo sapi
45	381.2	9.7	188413	2	AC145834	AC145834 Homo sapi

ALIGNMENTS

RESULT 1	AX528729	Sequence 15 from Patent WO0236780.	4150 bp	DNA	linear	PAT 21-NOV-2002
AX528729	LOCUS					
DEFINITION	Sequence 15 from Patent WO0236780.					
ACCESSION	AX528729					
VERSION	AX528729.1	GI:25172800				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
TITLE	Prouty, S.M., Zhang, L. and Stenn, K.S.					
JOURNAL	Stearoyl-coa desaturase gene promoter					
	Patent: WO 0236780-A 15 10-MAY-2002;					

1981 ATACAAAAAATTAGCTAAGTGTGGTGGCGCATGCTGTAAATCCACGACTACTTGGAGGGT 2040
2041 GAGCAGAGAAATTTCTTGAACCCGGAGGAGAGAGTTGCACTGAAGCGAGATCACACCA 2100
2041 GAGCAGAGAAATTTCTTGAACCCGGAGGAGAGAGTTGCACTGAAGCGAGATCACACCA 2100
2101 CTGCACTCCAGCCTTGGGGGAGAGAGCGAGACTTCTCTCTCAAAAAAACAACAAAGA 2160
2101 CTGCACTCCAGCCTTGGGGGAGAGAGCGAGACTTCTCTCTCAAAAAAACAACAAAGA 2160
2161 ATTAAGCAAAATTAGACATTTGACAGAGAACTCTGAAGGGGTTCAGACCACGACTACAGATTTC 2220
2161 ATTAAGCAAAATTAGACATTTGACAGAGAACTCTGAAGGGGTTCAGACCACGACTACAGATTTC 2220
2221 TGTGCCATATCCCAAGTACTTCTGAGGAGATCACTGGATGAGCTGTCCACATCTGAATCA 2280
2221 TGTGCCATATCCCAAGTACTTCTGAGGAGATCACTGGATGAGCTGTCCACATCTGAATCA 2280
2281 TCCAGTCTTGTTCAGAACTTTTCAACCGGACAGGGAGCCAGGACTGGAATGCACTCCT 2340
2281 TCCAGTCTTGTTCAGAACTTTTCAACCGGACAGGGAGCCAGGACTGGAATGCACTCCT 2340
2341 GGTCACTGGCCAGAGAGTTGGCCCTTGACCCCTGACACAGTGGCCCAAAAGAGAGCTGTT 2400
2341 GGTCACTGGCCAGAGAGTTGGCCCTTGACCCCTGACACAGTGGCCCAAAAGAGAGCTGTT 2400
2401 AGTCTACCTCCAGGAAATCCAGAGTCTTGTCTTCTCTGGGAAGTGAATCAATGGCGCAG 2460
2401 AGTCTACCTCCAGGAAATCCAGAGTCTTGTCTTCTCTGGGAAGTGAATCAATGGCGCAG 2460
2461 CACTCCGTAATTTCTCTCTCCAGGGGAAGATCTTAGGGCAGTATTTGGGAAAGACA 2520
2461 CACTCCGTAATTTCTCTCTCCAGGGGAAGATCTTAGGGCAGTATTTGGGAAAGACA 2520
2521 TGGGCATGGAAGACACCGGCTGAATCATAGCTGCTGCTTCTGAGACTCTCATGCTAA 2580
2521 TGGGCATGGAAGACACCGGCTGAATCATAGCTGCTGCTTCTGAGCTCTCATGCTAA 2580
2581 GGCTCTACAGACACGAAAGATGGGGGACAGGACAGATCAGTAGGGTTCAGAGCATC 2640
2581 GGCTCTACAGACACGAAAGATGGGGGACAGGACAGATCAGTAGGGTTCAGAGCATC 2640
2641 TCAGGACCCAGGGCAATATGTTCTCAGCAGGATTAAGAGCTTGGGCTCTCATATGGT 2700
2641 TCAGGACCCAGGGCAATATGTTCTCAGCAGGATTAAGAGCTTGGGCTCTCATATGGT 2700
2701 GTTCTGGGCTCAACTGCCAGCTCCGTCACCTTACTGCTGTGTGACCATGGGCAAGTTA 2760
2701 GTTCTGGGCTCAACTGCCAGCTCCGTCACCTTACTGCTGTGTGACCATGGGCAAGTTA 2760
2761 TTCATCTCTCCATATCTCTTCTCTCACTTTAAATGGAATTAATGGGTAATTAATTC 2820
2761 TTCATCTCTCCATATCTCTTCTCTCACTTTAAATGGAATTAATGGGTAATTAATTC 2820
2821 CAGGGTCACAGAGAGGCTTACAGAAACGATTTCTGTAATTCGCTTGCAGTAATTAATTC 2880
2821 CAGGGTCACAGAGAGGCTTACAGAAACGATTTCTGTAATTCGCTTGCAGTAATTAATTC 2880
2881 AATACCTGCCAGCTATTTCTTATTTCCATATCCAGGCCCTTTTCGCTGCTGCTGGTGAATA 2940
2881 AATACCTGCCAGCTATTTCTTATTTCCATATCCAGGCCCTTTTCGCTGCTGCTGGTGAATA 2940
2941 CACATGTCAGTGTTCCTGACGGTTTCCACAAAGAGATTTCCAAAATTAACAACCTGCCAG 3000
2941 CACATGTCAGTGTTCCTGACGGTTTCCACAAAGAGATTTCCAAAATTAACAACCTGCCAG 3000
3001 TCTGAAGAATCTCCAAAACATCCGACGATCTCTGAGCGCGGGCTTGGGATGGAC 3060
3001 TCTGAAGAATCTCCAAAACATCCGACGATCTCTGAGCGCGGGCTTGGGATGGAC 3060
3061 TGCCCGCCCGGGTCTCTGAACAGGATGCTGCGCGGACAGGACACACACAGCAGCCTG 3120
3061 TGCCCGCCCGGGTCTCTGAACAGGATGCTGCGCGGACAGGACACACACAGCAGCCTG 3120

3121 TGTCTGGCGCGAGTCCGCTGCGGTCCCGGTGAGCAGCGCCTGCTGCTGGCGGGC 3180
3121 TGTCTGGCGCGAGTCCGCTGCGGTCCCGGTGAGCAGCGCCTGCTGCTGGCGGGC 3180
3181 AGAGCCATTTGTCAGCGCGTACCGAGCCCGCGCTGCGCGGGAGGAGCGGGCT 3240
3181 AGAGCCATTTGTCAGCGCGTACCGAGCCCGCGCTGCGCGGGAGGAGCGGGCT 3240
3241 TCCCGCTTCCCAAGCTTCCAGATCTCTGGGTGGCTGCCACGTCTCCCTGCCACGCGCTG 3300
3241 TCCCGCTTCCCAAGCTTCCAGATCTCTGGGTGGCTGCCACGTCTCCCTGCCACGCGCTG 3300
3301 GGGGACCGGAAGACCGGAGCGAGATGTTAGTGTGGCGGCGCCCGCCAGGCTTCACCACT 3360
3301 GGGGACCGGAAGACCGGAGCGAGATGTTAGTGTGGCGGCGCCCGCCAGGCTTCACCACT 3360
3361 GTTTCCTGAGAACTTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGT 3420
3361 GTTTCCTGAGAACTTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGT 3420
3421 CCTGGCTAGGCTCCGCGCCCGGAGCGAGCTTGGGTGGCTGGGTGGGTGGGTGGGTGGGT 3480
3421 CCTGGCTAGGCTCCGCGCCCGGAGCGAGCTTGGGTGGCTGGGTGGGTGGGTGGGTGGGT 3480
3481 GCTCCCGACCGGGATGCGCGGCGAGAGCGCGCGGTGGGAAGAGAGTGGAGAGG 3540
3481 GCTCCCGACCGGGATGCGCGGCGAGAGCGCGCGGTGGGAAGAGAGTGGAGAGG 3540
3541 AGAAACAGAGGGAGGGGAGCGAGAGCTTGGCGGACAGAGGAAACAGCAGATTGCGCGA 3600
3541 AGAAACAGAGGGAGGGGAGCGAGAGCTTGGCGGACAGAGGAAACAGCAGATTGCGCGA 3600
3601 GCCAATGGCAACGCGAGCGAGCTGCGCAAAATTCCTTTCGGCCCAATCAGCAGCCGGA 3660
3601 GCCAATGGCAACGCGAGCGAGCTGCGCAAAATTCCTTTCGGCCCAATCAGCAGCCGGA 3660
3661 GTTTACAGAGCCTCATTTAGCATTTTCCAGAGCGAGGGGACAGGGCGAGAGCGGGTGG 3720
3661 GTTTACAGAGCCTCATTTAGCATTTTCCAGAGCGAGGGGACAGGGCGAGAGCGGGTGG 3720
3721 TGTGGTGTGGTGTGGCAGCATTCCTGGCGGCGCTGTGCGGTGCGCGGAGCTTCGGCC 3780
3721 TGTGGTGTGGTGTGGCAGCATTCCTGGCGGCGCTGTGCGGTGCGCGGAGCTTCGGCC 3780
3781 TCTGTCTCTCCCTCCCGCTTACCTTCCACGCGGACCGCCCGCGCAGTCAACTCC 3840
3781 TCTGTCTCTCCCTCCCGCTTACCTTCCACGCGGACCGCCCGCGCAGTCAACTCC 3840
3841 TCGCACTTTGCCCCCTGTTGGCAGCGGATAAAAGGGGGCTGAGGAAATACCGACACGGT 3900
3841 TCGCACTTTGCCCCCTGTTGGCAGCGGATAAAAGGGGGCTGAGGAAATACCGACACGGT 3900
3901 CACCCGTTGC 3910
3901 CACCCGTTGC 3910

RESULT 2
AF320307

LOCUS
DEFINITION

AF320307
VERSION

AF320307.1
KEYWORDS

ORGANISM
SOURCE

REFERENCE
AUTHORS

TITLE

4150 bp DNA linear PRI 06-DEC-2001
Homo sapiens stearyl-CoA desaturase (SCD) gene, exon 1 and intron 1 and partial cds.

GI:14150490

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 4150)

Zhang, L., Ge, L., Tran, T., Stenn, K. and Prouty, S.M.

Isolation and characterization of the human stearyl-CoA desaturase

gene promoter: requirement of a conserved CCAAT cis-element
Biochem. J. 357 (Pt 1), 183-193 (2001)
21308466
MEDLINE
PUBMED
11415448
REFERENCE
2 (bases 1 to 4150)
Zhang, L., Ge, L., Tran, T., Stenn, K. and Prouty, S.M.
Direct Submission
TITLE
Submitted (09-NOV-2000) Skin Biology TRC, Johnson and Johnson,
JOURNAL
CPW, 199 Grandview Road, Skillman, NJ 08558, USA

FEATURES
Location/Qualifiers
1..4150
source
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/tissue_type="scalp skin keratinocytes; liver; hair
follicles; brain"
3602..>4150
gene
/gene="SCD"
CAAT_signal
3602..3606
/gene="SCD"
TATA_signal
3669..3872
/gene="SCD"
TATA_signal
3924..3927
/gene="SCD"
mRNA
3958..>4129
/gene="SCD"
/product="stearoyl-CoA desaturase"
/note="3602"
3958..4129
exon
/gene="SCD"
/number=1
5'UTR
3958..4102
/gene="SCD"
CDS
4103..>4129
/gene="SCD"
/codon_start=1
product="stearoyl-CoA desaturase"
protein_id="AAK54510.1"
db_xref="GI:14150491"
translation="MPHLLQDD"
4130..>4150
intron
/gene="SCD"
/number=1

ORIGIN
Query Match 100.0%; Score 3910; DB 9; Length 4150;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGAATTCATCCATTTAAATCATACAAATTTATGCTTTTAGTATATTCACAGGTTGTC 60
DB 1 AGGAATTCATCCATTTAAATCATACAAATTTATGCTTTTAGTATATTCACAGGTTGTC 60
QY 61 ATCCATCAATCCATTTTAGAAGAGTTTATTTACTCCAAAATAAACCCTGCAATTCCTT 120
DB 61 ATCCATCAATCCATTTTAGAAGAGTTTATTTACTCCAAAATAAACCCTGCAATTCCTT 120
QY 121 AGCCATCACCCCCCAACATCTCCATCTCTCCATCCAGCCCTGGGCAACCAACATATAC 180
DB 121 AGCCATCACCCCCCAACATCTCCATCTCTCCATCCAGCCCTGGGCAACCAACATATAC 180
QY 181 TTTCTGTCTCTATAATTTGCCAATTCCTGGACATTTTCATATATAATGGAGCAACACAT 240
DB 181 TTTCTGTCTCTATAATTTGCCAATTCCTGGACATTTTCATATATAATGGAGCAACACAT 240
QY 241 GTGAGACTTTGTGACTGCTGCTTTCACTTAGCAATTCATTTTAAAGGCTCATTTATGTTA 300
DB 241 GTGAGACTTTGTGACTGCTGCTTTCACTTAGCAATTCATTTTAAAGGCTCATTTATGTTA 300
QY 301 CAGTACTTAGCAGTACTTCATCTTTTATCTCAAAAGTATCCACTGTGGGTAT 360
DB 301 CAGTACTTAGCAGTACTTCATCTTTTATCTCAAAAGTATTCACCTGTGGGTAT 360

QY 361 CCATATCATATATATATAGAGACAGGTTCTCACTCTGTGACCCAGGCTGGAGTGCAGTGGC 420
DB 361 CCATATCATATATATATAGAGACAGGTTCTCACTCTGTGACCCAGGCTGGAGTGCAGTGGC 420
QY 421 ACAATCATAGTCACTGTAACTCAAACTCCTGGGCTCAAGTGCATCTACTACTCAGCC 480
DB 421 ACAATCATAGTCACTGTAACTCAAACTCCTGGGCTCAAGTGCATCTACTACTCAGCC 480
QY 481 TCCAGAGTAGCTAGGACTACAGGCACACACAGGCATACCTGGCTAAATTTTTTTTAAAT 540
DB 481 TCCAGAGTAGCTAGGACTACAGGCACACACAGGCATACCTGGCTAAATTTTTTTTAAAT 540
QY 541 TTTCAATTTATGTATTCATTTTCTTTCTTTTGTGTTGTTGTTTGTAGATAGGCTCTC 600
DB 541 TTTCAATTTATGTATTCATTTTCTTTCTTTTGTGTTGTTGTTTGTAGATAGGCTCTC 600
QY 601 ACTTTGTTACCCAGGCTGGAGGCGAGTGGCATGGTGACAGCTGAGCAGCCTTGACTTCCT 660
DB 601 ACTTTGTTACCCAGGCTGGAGGCGAGTGGCATGGTGACAGCTGAGCAGCCTTGACTTCCT 660
QY 661 GGGCTCAAGTGCATCTCCTGCTCAGCCTCCCAAGTAGCTGGAGCTACAAAACAGCTGCA 720
DB 661 GGGCTCAAGTGCATCTCCTGCTCAGCCTCCCAAGTAGCTGGAGCTACAAAACAGCTGCA 720
QY 721 CCATGCTGGCTGATATATTTTCTTGAACACAGGCTATCACTCTGTTGCCAGGCTGGA 780
DB 721 CCATGCTGGCTGATATATTTTCTTGAACACAGGCTATCACTCTGTTGCCAGGCTGGA 780
QY 781 GTACAGTGGCTGAATTAATAGTCACTGAGCCTCCCTCCTGGGCTCAAGCAATCCGCTG 840
DB 781 GTACAGTGGCTGAATTAATAGTCACTGAGCCTCCCTCCTGGGCTCAAGCAATCCGCTG 840
QY 841 GCCTCAGCATCTGAGTAGCTGGGACTACAGCTTGTGCCACCGCCAGCCAGCTAGATTTT 900
DB 841 GCCTCAGCATCTGAGTAGCTGGGACTACAGCTTGTGCCACCGCCAGCCAGCTAGATTTT 900
QY 901 AAAAAATGATTTTGGTATAGAGGAGTCTTGTCTATGTTGCTCAGGCTGTATTTTATTTG 960
DB 901 AAAAAATGATTTTGGTATAGAGGAGTCTTGTCTATGTTGCTCAGGCTGTATTTTATTTG 960
QY 961 TTGAGACAAGGCTCACTATGTTGCCATGATCCGCCACCTCCACTTCCCAAAGTGCTCA 1020
DB 961 TTGAGACAAGGCTCACTATGTTGCCATGATCCGCCACCTCCACTTCCCAAAGTGCTCA 1020
QY 1021 TCTTATCTGTTTCATTTAGTCACTTGACAGACATTTAGGTTGTTTCCACTTTTGGACATTA 1080
DB 1021 TCTTATCTGTTTCATTTAGTCACTTGACAGACATTTAGGTTGTTTCCACTTTTGGACATTA 1080
QY 1081 TGAATTAATCTCCAGTGAATTTCAATGATACATTTGTGTGGGCATATGTTTTCATTTCT 1140
DB 1081 TGAATTAATCTCCAGTGAATTTCAATGATACATTTGTGTGGGCATATGTTTTCATTTCT 1140
QY 1141 GTTGGGTTTATCTAGAGTGGAAATTTGCTGATCCCGGGTAAATTTTGCACAGGCAGAG 1200
DB 1141 GTTGGGTTTATCTAGAGTGGAAATTTGCTGATCCCGGGTAAATTTTGCACAGGCAGAG 1200
QY 1201 TTCAGGGGAAGAAAACCTTGGGAAAATGAGCATGTTTAGAAAATCAGCAAGAGTGCAAGG 1260
DB 1201 TTCAGGGGAAGAAAACCTTGGGAAAATGAGCATGTTTAGAAAATCAGCAAGAGTGCAAGG 1260
QY 1261 GTTTTTCGGAGTTTATTTTATATCTGTTGACAAATGTCAGTTTTCATGAAGATACAAG 1320
DB 1261 GTTTTTCGGAGTTTATTTTATTTCTGTTGACAAATGTCAGTTTTCATGAAGATACAAG 1320
QY 1321 TTATACTAAGTGAGAGTGAATTAAGGCTGGAATAGGCGTTTCAGAGTAAATCATGA 1380
DB 1321 TTATACTAAGTGAGAGTGAATTAAGGCTGGAATAGGCGTTTCAGAGTAAATCATGA 1380
QY 1381 AGCACTTTGTAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATAATAAAAAATCAAA 1440
DB 1381 AGCACTTTGTAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATAATAAAAAATCAAA 1440

QY	1441	TTTTTTTTTTTTTTTTTTGAGAAAGAGTCTTGCTCTTTTCACCCCTGGCTGAGGGCAGTGGTG	1500
Db	1441	TTTTTTTTTTTTTTTTTTGAGAAAGAGTCTTGCTCTTTTCACCCCTGGCTGAGGGCAGTGGTG	1500
QY	1501	TGATCTCAGCTCAGTGCACAACTTTCCGCTCCCGGGTTCAAGCAATCTCCTGCTCAGCCT	1560
Db	1501	TGATCTCAGCTCAGTGCACAACTTTCCGCTCCCGGGTTCAAGCAATCTCCTGCTCAGCCT	1560
QY	1561	CCCAAGTAGCTGGGACTACAGGCACTTCCCAACCATGCCCCAGCTGATTTTGTATTTTAG	1620
Db	1561	CCCAAGTAGCTGGGACTACAGGCACTTCCCAACCATGCCCCAGCTGATTTTGTATTTTAG	1620
QY	1621	TAGAGATGGGATTTACATTTGTGCGCAAGCTGGTCTCAAACTTTTCTGTCATATTTG	1680
Db	1621	TAGAGATGGGATTTACATTTGTGCGCAAGCTGGTCTCAAACTTTTCTGTCATATTTG	1680
QY	1681	TTGTAACATTAATTTCTTTTCTGCTGAGGTAGGGCCCCCAGACCAAAAAATAAATCTTTAG	1740
Db	1681	TTGTAACATTAATTTCTTTTCTGCTGAGGTAGGGCCCCCAGACCAAAAAATAAATCTTTAG	1740
QY	1741	AATCCAAATCAGTGTGTTGGTTTGACCACTGTCTACCTTGAGAACCAAGTGTGACAGGGC	1800
Db	1741	AATCCAAATCAGTGTGTTGGTTTGACCACTGTCTACCTTGAGAACCAAGTGTGACAGGGC	1800
QY	1801	CTCAGGAGTAGAGTGATCTCTGCTCGAAAGAGAAATAGAAATGAATATTTCTCCGGGC	1860
Db	1801	CTCAGGAGTAGAGTGATCTCTGCTCGAAAGAGAAATAGAAATGAATATTTCTCCGGGC	1860
QY	1861	AGGCGTGGTGGCTCATGCTGTAATCCAGCACCTTTGGGAGGCCAAGGCATGTGGATCAC	1920
Db	1861	AGGCGTGGTGGCTCATGCTGTAATCCAGCACCTTTGGGAGGCCAAGGCATGTGGATCAC	1920
QY	1921	CTGAGGTTCAGGAGTTCAAAACAGCCTGGCCAAATGCTGTAATCCAGCTACTTTGGAGGGT	1980
Db	1921	CTGAGGTTCAGGAGTTCAAAACAGCCTGGCCAAATGCTGTAATCCAGCTACTTTGGAGGGT	1980
QY	1981	ATACAAAATAATAGCTAAGTGTGGTGGCGCATGCTGTAATCCAGCTACTTTGGAGGGT	2040
Db	1981	ATACAAAATAATAGCTAAGTGTGGTGGCGCATGCTGTAATCCAGCTACTTTGGAGGGT	2040
QY	2041	GAGCAGGAGAAATTTCTTGAACCCCGGAGGAGAGGTTGCAAGTGAAGCGATCACACCA	2100
Db	2041	GAGCAGGAGAAATTTCTTGAACCCCGGAGGAGAGGTTGCAAGTGAAGCGATCACACCA	2100
QY	2101	CTGCACCTCCAGCTGGGGAGAGAGCGAGCTTCTCTCAAAAAACAAAAACAAAAA	2160
Db	2101	CTGCACCTCCAGCTGGGGAGAGAGCGAGCTTCTCTCAAAAAACAAAAACAAAAA	2160
QY	2161	ATTAAGCAAAATTAGACATTGCAGAGAAACCTGAAGGGGGTTCAGACCACTGACAGATTTC	2220
Db	2161	ATTAAGCAAAATTAGACATTGCAGAGAAACCTGAAGGGGGTTCAGACCACTGACAGATTTC	2220
QY	2221	TGTGCCACATGCCAAGTACTTCTGAGGCATGACTGGATGAGCTGTCCACATCTGAAATCA	2280
Db	2221	TGTGCCACATGCCAAGTACTTCTGAGGCATGACTGGATGAGCTGTCCACATCTGAAATCA	2280
QY	2281	TCCAGTCTTGTTCAGAACTTTCACCGGACAGGGACCGAGGCTGGAATGCGATCTCCT	2340
Db	2281	TCCAGTCTTGTTCAGAACTTTCACCGGACAGGGACCGAGGCTGGAATGCGATCTCCT	2340
QY	2341	GGTCACTGGCCAGAGAGTTGGCCCTTGAACCTGAGACCAAGTGGGCCAAACAAAGAGCTGCTT	2400
Db	2341	GGTCACTGGCCAGAGAGTTGGCCCTTGAACCTGAGACCAAGTGGGCCAAACAAAGAGCTGCTT	2400
QY	2401	AGTCTACCTCCAGGAAATCCAGGTGCTTGTCTTCTCTGGGAAAGTGAATCATTTGGCGCAG	2460
Db	2401	AGTCTACCTCCAGGAAATCCAGGTGCTTGTCTTCTCTGGGAAAGTGAATCATTTGGCGCAG	2460
QY	2461	CACCTCCGTAATTTCTCTCTTCTCCAGGGAGAGTCTTAGGGCAGTATTTGGAAAGACA	2520
Db	2461	CACCTCCGTAATTTCTCTCTTCTCCAGGGAGAGTCTTAGGGCAGTATTTGGAAAGACA	2520
QY	2521	TGGGCATGGAAGACACCGGGTGAATGCATAGCCTGCTGGTTCTGAGCTCTCATGGTAA	2580

Db	2521	TGGGCATGGAAGACACCGGGTGAATGCATAGCCTGCTGGTTCTGAGCTCTCATGGTAA	2580
QY	2581	GGCTCTCTACACACACGGAAGATGGGGCACAGGACAGATCAGTAGGGTCAGAGCATC	2640
Db	2581	GGCTCTCTACACACACGGAAGATGGGGCACAGGACAGATCAGTAGGGTCAGAGCATC	2640
QY	2641	TCAGGGACCGAGGGCAATATATGCTCTGAGCAGGGAATTAAGAGCTTGGGCTCTCATATGGT	2700
Db	2641	TCAGGGACCGAGGGCAATATATGCTCTGAGCAGGGAATTAAGAGCTTGGGCTCTCATATGGT	2700
QY	2701	GTTTTCTGGGCTCAACTGCCAGCTCGTCACTTACTGTTGCTGTGACCATGGGCAAGTTA	2760
Db	2701	GTTTTCTGGGCTCAACTGCCAGCTCGTCACTTACTGTTGCTGTGACCATGGGCAAGTTA	2760
QY	2761	TTCCATCTCTCATATCTCTTTCTCACTTTTAAATGGAATAATGGGGTACCCACCTCC	2820
Db	2761	TTCCATCTCTCATATCTCTTTCTCACTTTTAAATGGAATAATGGGGTACCCACCTCC	2820
QY	2821	CAGGTCACAGAGGGCTTACAGAAAACGATTTCTGTGAATTTGGCTTGCAGTAATAATTC	2880
Db	2821	CAGGTCACAGAGGGCTTACAGAAAACGATTTCTGTGAATTTGGCTTGCAGTAATAATTC	2880
QY	2881	AATACCTGCCAGCTATTTCTTATTTCCACATCCAAAGCCCTTTGCGCTGCTGGGTGAAA	2940
Db	2881	AATACCTGCCAGCTATTTCTTATTTCCACATCCAAAGCCCTTTGCGCTGCTGGGTGAAA	2940
QY	2941	CACATGTCAAGTGTTCCTGACGGTTTCCAAAAAGATTTCCAAAAATTTACAACTTGCAC	3000
Db	2941	CACATGTCAAGTGTTCCTGACGGTTTCCAAAAAGATTTCCAAAAATTTACAACTTGCAC	3000
QY	3001	TCGTAGAATCTCAAAAACATTCGCGACGATCTCGAGCGCGGGCTTGGGGATGGGAC	3060
Db	3001	TCGTAGAATCTCAAAAACATTCGCGACGATCTCGAGCGCGGGCTTGGGGATGGGAC	3060
QY	3061	TGCCCGCCCGGGTCTCTGAAAGAGTGCCTGCGCGAGGACACACACAGCAGCCTG	3120
Db	3061	TGCCCGCCCGGGTCTCTGAAAGAGTGCCTGCGCGAGGACACACACAGCAGCCTG	3120
QY	3121	TGTGTGCGCGCGAGTCCGGTCCCGGGTTCAGCAGCGCTGCTGGTGGCGGGGC	3180
Db	3121	TGTGTGCGCGCGAGTCCGGTCCCGGGTTCAGCAGCGCTGCTGGTGGCGGGGC	3180
QY	3181	AGAGCCATTTGTCGAGGCTTACGAGCCCCCGCTGCGCGGAGGAGGAGGCGGGCT	3240
Db	3181	AGAGCCATTTGTCGAGGCTTACGAGCCCCCGCTGCGCGGAGGAGGAGGCGGGCT	3240
QY	3241	TCCCGGCTCCCAAGCTCCAGATCCTGGGGTGGCTGCCAGTCTCCCTGCCACGCGCTG	3300
Db	3241	TCCCGGCTCCCAAGCTCCAGATCCTGGGGTGGCTGCCAGTCTCCCTGCCACGCGCTG	3300
QY	3301	GGGGGACGGGAACACGGGACGGAGATGTTAGTGGTGGCGGCCCGCCGAGGGTTCCACACT	3360
Db	3301	GGGGGACGGGAACACGGGACGGAGATGTTAGTGGTGGCGGCCCGCCGAGGGTTCCACACT	3360
QY	3361	GTTTTCTGAGAAATTTCCCAAGTGGCCACCCAGTCTTCTCGTGTGCCGAGGGCGGT	3420
Db	3361	GTTTTCTGAGAAATTTCCCAAGTGGCCACCCAGTCTTCTCGTGTGCCGAGGGCGGT	3420
QY	3421	CCTGGGCTAGGCTCCCGGCCCGCCAGCCCCCAACCGGGTCCCGAGCCCTTCCAGAGAAA	3480
Db	3421	CCTGGGCTAGGCTCCCGGCCCGCCAGCCCCCAACCGGGTCCCGAGCCCTTCCAGAGAAA	3480
QY	3481	GCTCCCGACCGGGGATGCGCGGACAGGGCCCCAGCGGGGTGGAAGAGAGCTGAGAAG	3540
Db	3481	GCTCCCGACCGGGGATGCGCGGACAGGGCCCCAGCGGGGTGGAAGAGAGCTGAGAAG	3540
QY	3541	AGAAACAGAGGGGAGGGGAGCGAGAGCTGGCGGAGAGGAGAACAGCAGATTGCCCGGA	3600
Db	3541	AGAAACAGAGGGGAGGGGAGCGAGAGCTGGCGGAGAGGAGAACAGCAGATTGCCCGGA	3600
QY	3601	GCCATGGCAACCGGAGGACGAGGTGGCACCAAAATTTCCCTTGGCCCAATGACGAGCCGGA	3660

```

Db      3601  GCCAATGGCAACGCGCAGGAGGTGGCAACAAATTCCTTCGGCCCAATGACGAGCGCGA 3660
QY      3661  GTTTACAGAGCCTCATTTAGCATTTCCCGAGAGGCGAGGGCAGAGGCGCGGTGG 3720
Db      3661  GTTTACAGAGCCTCATTTAGCATTTCCCGAGAGGCGAGGGCAGAGGCGCGGTGG 3720
QY      3721  TGTGGTGTGGTGTGGCAGCATCCCGGGCGCCCTGCTGGCGTGGCGCGGAGCCTCGGCC 3780
Db      3721  TGTGGTGTGGTGTGGCAGCATCCCGGGCGCCCTGCTGGCGTGGCGCGGAGCCTCGGCC 3780
QY      3781  TCTGTCTCTCCCTCCCGCTTACTCGACGCGGAGCCGCGCGGAGCCGCGCGCACTCACTCC 3840
Db      3781  TCTGTCTCTCCCTCCCGCTTACTCGACGCGGAGCCGCGCGGAGCCGCGCGCACTCACTCC 3840
QY      3841  TCGCACTTTGCCCTGCTTCGCGAGCGATAAAGGGGGCTGAGGAATACCGGACACGGT 3900
Db      3841  TCGCACTTTGCCCTGCTTCGCGAGCGATAAAGGGGGCTGAGGAATACCGGACACGGT 3900
QY      3901  CACCCGTGC 3910
Db      3901  CACCCGTGC 3910

RESULT 3
AC018783/c
LOCUS   Homo sapiens clone RP11-2M13, WORKING DRAFT SEQUENCE, 12 unordered
DEFINITION
ACCESSION AC018783
VERSION   AC018783.3 GI:7223991
KEYWORDS  HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 159351)
AUTHORS   Birren,B., Linton,L., Nussbaum,C., Lander,B., Abrahams,H., Allen,N.,
           Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A.,
           Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
           DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenesor,J.,
           Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
           Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
           Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
           Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
           Macdonald,P., Marguis,N., McEwan,P., McGurk,A., McKernan,K.,
           Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H.,
           O'Connor,T., O'Donnell,P., Olivier,T.M., Peterson,K., Pierre,N.,
           Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,
           Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
           Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
           Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.
           and Zody,M.
DIRECT SUBMISSION
Submitted (19-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 12, 2000 this sequence version replaced gi:6692332.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2667
Center clone name: 2_M13
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 151339 bases at least Q40
Consensus quality: 155177 bases at least Q30
Consensus quality: 157011 bases at least Q20
Insert size: 167000; agarose-fp
Insert size: 158251; sum-of-ctngs
Quality coverage: 4.9 in Q20 bases; agarose-fp
Quality coverage: 5.2 in Q20 bases; sum-of-ctngs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2297: contig of 2297 bp in length
* 2298 2397: gap of 100 bp
* 2398 5719: contig of 3322 bp in length
* 5720 5819: gap of 100 bp
* 5820 10149: contig of 4330 bp in length
* 10150 10249: gap of 100 bp
* 10250 16496: contig of 6247 bp in length
* 16497 16596: gap of 100 bp
* 16597 21823: contig of 5227 bp in length
* 21824 21923: gap of 100 bp
* 21924 25747: contig of 3824 bp in length
* 25748 25847: gap of 100 bp
* 25848 31971: contig of 6124 bp in length
* 31972 32071: gap of 100 bp
* 32072 44449: contig of 12378 bp in length
* 44450 44549: gap of 100 bp
* 44550 64978: contig of 20429 bp in length
* 64979 65079: gap of 100 bp
* 65079 91491: contig of 26413 bp in length
* 91492 121937: contig of 30346 bp in length
* 121938 122038: gap of 100 bp
* 122038 159351: contig of 37314 bp in length.
Location/Qualifiers
1..159351
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-2M13"
/clone_lib="RPC1-11 Human Male BAC"
1..2297
/note="assembly_fragment"
2398..5719
/note="assembly_fragment"
5820..10149
/note="assembly_fragment"
10250..16496
/note="assembly_fragment"
16597..21823
/note="assembly_fragment"
21924..25747
/note="assembly_fragment"
clone_end:SP6
vector_side:right
25848..31971
/note="assembly_fragment"
clone_end:17
vector_side:left
32072..44449
/note="assembly_fragment"
44550..64978
/note="assembly_fragment"
65079..159351
/note="assembly_fragment"

```


QY	2043	GGCAGGAGAAATTTCTTGAACCCGGGAGGACAGAGTTTGCAGTGAAGCGACATCACACCACT	2107
Db	57350	GGCAGGAGAAATTTCTTGAACCCGGGAGGACAGAGTTTGCAGTGAAGCGACATCACACCACT	57291
QY	2103	GCACCTCCAGCCTGGGGAGAGCGGAGACTTCTCTCAAAAAACAATAAAACAAAAGAAAT	2162
Db	57290	GCACCTCCAGCCTGGGGAGAGCGGAGACTTCTCTCAAAAAACAATAAAACAAAAGAAAT	57231
QY	2163	TAAGCAAAATAGACATTTGCAGAGAGAACTGAGGGGTGCAGACACGTCAGATTCTTG	2222
Db	57230	TAGACAAATAGACATTTGCAGAGAGAACTGAGGGGTGCAGACACGTCAGATTCTTG	57171
QY	2223	TGCCACATGCAAGTACTTCTGTAGGCATGACTGGATGAGCTGTCCACATCTCGAAATCATC	2282
Db	57170	TGCCACATGCAAGTACTTCTGTAGGCATGACTGGATGAGCTGTCCACATCTCGAAATCATC	57111
QY	2283	CAGTCTTTGTTTCAGAACTTTTCACACCGGACAGGAGCGAGGACTGGAATGCGATCTCTGG	2342
Db	57110	CAGTCTTTGTTTCAGAACTTTTCACACCGGACAGGAGCGAGGACTGGAATGCGATCTCTGG	57051
QY	2343	TCACTGGCCAGAGAGTTGGCCTTGACCTTGACACCAAGTGGCCCAACAAAGGAGCTGCTTAG	2402
Db	57050	TCACTGGCCAGAGAGTTGGCC-TGACCTTGACACCAAGTGGCCCAACAAAGGAGCTGCTTAG	56992
QY	2403	TCTACCTCCAGGAAATCCAGGTCTTGTCTTCTCTGGGAAATGAATCATTTGGCGCGACGA	2462
Db	56991	TCTACCTCCAGGAAATCCAGGTCTTGTCTTCTCTGGGAAATGAATCATTTGGCGCGACGA	56932
QY	2463	CTCCGTATTTCTCTCTTCCACAGGGAAGATCTTAGGGCAGTATTTGGGAAAGACATG	2522
Db	56931	CTCCGTATTTCTCTCTTCCACAGGGAAGATCTTAGGGCAGTATTTGGGAAAGACATG	56872
QY	2523	GGCATGGAAGGACACCGGTGAATGCATAGCTGCTGCTGTTCTGAGCTCTCATGTTAAGG	2582
Db	56871	GGCATGGAAGGACACCGGTGNAATGCATAGCTGCTGCTGTTCTGAGCTCTCATGTTAAGG	56812
QY	2583	CTCTCTACAGACACGGAAGAATGGGGGCACAGGACAGATCAGTAGGGTCAGAGCATCTC	2642
Db	56811	CTCTCTACAGACACGGAAGAATGGGGGCACAGGACAGATCAGTAGGGTCAGAGCATCTC	56752
QY	2643	AGGGACCCAGGGCATATGGTCTTGACAGGGATTAAGCTTTGGGCTCTCATATGGTGT	2702
Db	56751	AGGGACCCAGGGCATATGGTCTTGACAGGGATTAAGCTTTGGGCTCTCATATGGTGT	56692
QY	2703	TTCTGGGCTCAACTGCCAGCTCCGCTCACTTACTGGTTGCTGTGACCATGGGCAAGTATT	2762
Db	56691	TTCTGGGCTCAACTGCCAGCTCCATCACTTACTGGTTGCTGTGACCATGGGCAAGTATT	56632
QY	2763	CCATCTCTCCATATCTTCTTCTCAATTTAAATGGAATAATGGGTACCCACCTCCCA	2822
Db	56631	CCATCTCTCCATATCTTCTTCTCAATTTAAATGGAATAATGGGTACCCACCTCCCA	56572
QY	2823	GGGTACAGAGAGCTTTACAGAAAACGATTTCTGTGNAATTTGGCTTGCAGTAAATTCAA	2882
Db	56571	GGGTACAGAGAGCTTTACAGAAAACGATTTCTGTGNAATTTGGCTTGCAGTAAATTCAA	56512
QY	2883	TACCTGCCAGCTATCTTTATTTCCACATCCAAAGCCCTTTGCGCTGCTGCTGGGTGAACA	2942
Db	56511	TACCTGCCAGCTATCTTTATTTCCACATCCAAAGCCCTTTGCGCTGCTGCTGGGTGAACA	56452
QY	2943	CATCTAGTGTTCCTGCACCGGTTTCCACAAAGAGATTTCCAAATTTACAACCTGCCAGTC	3002
Db	56451	CATCTAGTGTTCCTGCACCGGTTTCCACAAAGAGATTTCCAAATTTACAACCTGCCAGTC	56392
QY	3003	TGAAGATCTCCAAAAATCCCGCACGATCTCTGGAGGCGCGGCTTGGGGATGGGACTG	3062
Db	56391	TGAAGATCTCCAAAAATCCCGCACGATCTCTGGAGGCGCGGCTTGGGGATGGGACTG	56332
QY	3063	CCGCGCCGGGTCTGTAAACAGGATGCGTGGCGCAGCACACACACGAGCGAGCTGTG	3122
Db	56331	CCGCGCCGGGTCTGTAAACAGGATGCGTGGCGCAGCACACACACGAGCGAGCTGTG	56272
QY	3123	TGTGCGCCGAGATTCGCGTGCGGTCCCGGTGAGCAGCGCGTGGCTGTGGCGCGGGCAG	3182

Db	56271	TGTGCGCCGGAGTCCGGTGCGGTCCCGGTGAGCAGCGTGGTGGTGGCGGGCAG	56212
QY	3183	AGCCATTGTTCCGACGGGTACCGAGCCCCCGCGCTGCCCGGAGGAGCGGGGCTTC	3242
Db	56211	AGCCATTGTTCCGACGGGTACCGAGCCCCCGCGCTGCCCGGAGGAGCGGGGCTTC	56152
QY	3243	CCGCGT-CCCCAAGCTCCAGATCTCGGGTGGCTGCCACGCTCCCTGCCACGCGCCTGG	3301
Db	56151	CCGCGTCCCCCAAGCTCCAGATCTCGGGTGGCTGCCACGCTCCCTGCCACGCGCCTGG	56092
QY	3302	GGGGA CGGGAAGACGGACCGAGATGTTAGTGTGGCGCCCCCGAGGGTTCCACTG	3361
Db	56091	GGGGA CGGGAAGACGGACCGAGATGTTAGTGTGGCGCCCCCGAGGGTTCCACTG	56032
QY	3362	TTTCTCGAGAAACTTCCCCAGTGCCACCCACCGCTTCTCCGTGTCGCCGAGGCGCGTC	3421
Db	56031	TTTCTCGAGAAACTTCCCCAGTGCCACCCACCGCTTCTCCGTGTCGCCGAGGCGCGTC	55972
QY	3422	CTGGGCTAGGCTCCGCGOCCAGCCCCAAACCGGGTCCCGAGCCCTTCCAGAGAGAAAG	3481
Db	55971	CTGGGCTAGGCTCCGCGOCCAGCCCCAAACCGGGTCCCGAGCCCTTCCAGAGAGAAAG	55912
QY	3482	CTCCCGACGCGGGATCCGCGGCAGAGGCCCGCGGGGTGGGAAGAGAGCTGAGAGGA	3541
Db	55911	CTCCCGACGCGGGATCCGCGGCAGAGGCCCGCGGGGTGGGAAGAGAGCTGAGAGGA	55852
QY	3542	GAACACAGAGGGAGGGGAGCCGAGGAGCTGGCGGCAGAGGGACACAGATTCGCGCGAG	3601
Db	55851	GAACACAGAGGGAGGGGAGCCGAGGAGCTGGCGGCAGAGGGACACAGATTCGCGCGAG	55792
QY	3602	CCAATGGCAACGGCGAGGAGGAGGTGGCAACCAATTCCCTTCGGCCCAATGACGAGCCGGAG	3661
Db	55791	CCAATGGCAACGGCGAGGAGGAGGTGGCAACCAATTCCCTTCGGCCCAATGACGAGCCGGAG	55732
QY	3662	TTTACAGAGGCTCATTTAGCATTTCCCGAGAGGCGGGGACAGGGCGAGGCGGGTGGT	3721
Db	55731	TTTACAGAGGCTCATTTAGCATTTCCCGAGAGGCGGGGACAGGGCGAGGCGGGTGGT	55672
QY	3722	GTGGTGTCCGTTCCGGCAGCATCCCCGGCGCCCTGCTGCGGTGCGCGGAGCCTCGGCCCT	3781
Db	55671	GTGGTGTCCGTTCCGGCAGCATCCCCGGCGCCCTGCTGCGGTGCGCGGAGCCTCGGCCCT	55612
QY	3782	CTGTCTCTCCCTTCCCGCCCTTACCTTCGAGCGGGACCGCCCGCGCAGTCAACTCT	3841
Db	55611	CTGTCTCTCCCTTCCCGCCCTTACCTTCGAGCGGGACCGCCCGCGCAGTCAACTCT	55552
QY	3842	CGCACTTTGCCCTTGTTCGAGCGGATAAAGGGGGCTGAGGAATAACCGGACACGGTC	3901
Db	55551	CGCACTTTGCCCTTGTTCGAGCGGATAAAGGGGGCTGAGGAATAACCGGACACGGTC	55492
QY	3902	ACCGTTGCG	3910
Db	55491	ACCGTTGCG	55483

RESULT 4	
LOCUS	AL139819
DEFINITION	Human DNA sequence from clone RP11-34D15 on chromosome 10, complete sequence.
ACCESSION	AL139819
VERSION	AL139819.8
KEYWORDS	HTG.
SOURCE	GI:15705162
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 115278)
AUTHORS	Tracey, A.
TITLE	Direct Submision
JOURNAL	Submitted (18-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humonerv@sanger.ac.uk.Closure

requests: clonerquest@sanger.ac.uk
On Sep 19, 2001 this sequence version replaced gi:14586033.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP. Information on the WORMPEP database can be found at

database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone configs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr10> RP11-34D15 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone Rp11-34D15. It may be shorter because we sequence overlapping sections only once, except for a short overlap. The true right end of clone Rp11-34D15 is at 115278 in this sequence. The true left end of clone Rp11-34D15 is at 56886 in this sequence. The true right end of clone Rp11-316M21 is at 2000 in this sequence.

```

in this sequence.
Location/Qualifiers
1. 115278
   /organism="Homo sapiens"
   /mol_type="genomic"
   /db_xref="taxon:9606"
   /chromosome="10"
   /clone="RP11-34D15"
   /clone_lib="RP11-34D15"

```

ORIGIN

Query Match	99.0%;	Score 3869.8;	DB 9;	Length 115276;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 3904;	Conservative 0;	Mismatches 2;	Indels 3;	Gaps 3;
QY	3	GAATTCATCCATTTAAATCATACAATTTAAATGCCTTTTAGTATATATTCACAGGTGTGCAT	62	
Db	64592	GAATTCATCCATTTAAATCATACAATTTAAATGCCTTTTAGTATATATTCACAGGTGTGCAT	64651	
QY	63	CCATCAAAATCCATTTTAGAACAGTTTATTATCTCCAAAAATAAACCCGTGCATCTCCTTAG	122	
Db	64652	CCATCAAAATCCATTTTAGAACAGTTTATTATCTCCAAAAATAAACCCGTGCATCTCCTTAG	64711	
QY	123	CCATCAACCCCCCAACATCCTTCATCTCTCCAGAGCCCTGGGCAACACCAATCTACTTT	182	
Db	64712	CCATCAACCCCCCAACATCCTTCATCTCTCCAGAGCCCTGGGCAACACCAATCTACTTT	64771	
QY	183	TCGTCTCTATAAAATTGGCAATTCGTGGACATTTTCATATAAATGAAGCAACCAACATGT	242	
Db	64772	TCGTCTCTATAAAATTGGCAATTCGTGGACATTTTCATATAAATGAAGCAACCAACATGT	64831	
QY	243	GAGACTTTGTGACTGGCTGCTTTCATCTTAGCAATCTCTATTTTAAAGGCTCATTTATGTACA	302	
Db	64832	GAGACTTTGTGACTGGCTGCTTTCATCTTAGCAATCTCTATTTTAAAGGCTCATTTATGTACA	64891	
QY	303	GTACTTAGCAGTACTTCATTCCTTTTTTATCTCAAAATGGTATTCACATGTGTGGGTATCC	362	
Db	64892	GTACTTAGCAGTACTTCATTCCTTTTTTATCTCAAAATGGTATTCACATGTGTGGGTATCC	64951	

|||||
66032 TTTT|TTTTTT|TTT|GAGAAA|GAGTCT|TGCTCTTT|CACCCCTGGCTGGAGGGCAGTGGTGTG 66091
QY 1503 ATCT|CAGCT|CACTG|CAACTTT|TGGCTCC|CGGTT|CAAGAA|TATCTCT|GCTTTCAGCCTCC 1562
Db 66092 ATCT|CAGCT|CACTG|CAACTTT|TGGCTCC|CGGTT|CAAGAA|TATCTCT|GCTTTCAGCCTCC 66151
QY 1563 CAAGT|AGCTGG|AGCTAC|AGGCACTTCC|CACATG|CCAGCTGATTTT|GTATTTTAGTA 1622
Db 66152 CAAGT|AGCTGG|AGCTAC|AGGCACTTCC|CACATG|CCAGCTGATTTT|GTATTTTAGTA 66211
QY 1623 GAGAT|GGATTT|CACTTT|TGTGG|CAAGCTGGTCT|CAAACTTTT|TGTGTCATATTTGTT 1682
Db 66212 GAGAT|GGATTT|CACTTT|TGTGG|CAAGCTGGTCT|CAAACTTTT|TGTGTCATATTTGTT 66271
QY 1683 GTA|ACTATTT|GTTCTTTT|TGTG|AGGAGTGGG|CCCCAG|ACCAAAA|AAATATAATCTTAGAA 1742
Db 66272 GTA|ACTATTT|GTTCTTTT|TGTG|AGGAGTGGG|CCCCAG|ACCAAAA|AAATATAATCTTAGAA 66331
QY 1743 TCCAAAT|CAGTGT|GTTGGT|TTGACCA|CTGTCA|CTTGAGAACCA|CAGTGTGAC|CAGGCT 1802
Db 66332 TCCAAAT|CAGTGT|GTTGGT|TTGACCA|CTGTCA|CTTGAGAACCA|CAGTGTGAC|CAGGCT 1862
QY 1803 CAGAGT|PAGAGT|GATCTCT|GTCTCG|AAAGAGAAATAGAA|TATAATTTCTCGG|GCCAG 1862
Db 66392 CAGAGT|PAGAGT|GATCTCT|GTCTCG|AAAGAGAAATAGAA|TATAATTTCTCGG|GCCAG 66451
QY 1863 GCGT|GTG|GCTCAT|GCTGTAT|TCCAG|CACTTTGG|AGG|CCAAAGGCATGTG|GATCACCT 1922
Db 66452 GCGT|GTG|GCTCAT|GCTGTAT|TCCAG|CACTTTGG|AGG|CCAAAGGCATGTG|GATCACCT 66511
QY 1923 GAGGT|CAGAGT|TCAAAA|CCAGCTTGG|CCAA|CATGTGTGAA|CCCCGTCTCTA|TAAAT 1982
Db 66512 GAGGT|CAGAGT|TCAAAA|CCAGCTTGG|CCAA|CATGTGTGAA|CCCCGTCTCTA|TAAAT 66571
QY 1983 A|CAAAAA|TTAGCTA|AGTGTGGT|GGCATG|CTGTAAT|TCCAGTACTT|GCGAGG|GTA 2042
Db 66572 A|CAAAAA|TTAGCTA|AGTGTGGT|GGCATG|CTGTAAT|TCCAGTACTT|GCGAGG|GTA 66631
QY 2043 GGCAG|GAGATTT|CTTGA|ACCCGG|AGG|CAGAGTGTG|CATGATGA|ACGAGAT|CACACCACT 2102
Db 66632 GGCAG|GAGATTT|CTTGA|ACCCGG|AGG|CAGAGTGTG|CATGATGA|ACGAGAT|CACACCACT 66691
QY 2103 GCAC|TCCAGCTTGGGG|GAGAG|CAGACTTCT|CTCAAAAA|AAACAAAA|CAAAAGAAAT 2162
Db 66692 GCAC|TCCAGCTTGGGG|GAGAG|CAGACTTCT|CTCAAAAA|AAACAAAA|CAAAAGAAAT 66751
QY 2163 TAAGCA|AATTAG|CAATTC|CAGAGAA|CCTGAA|GGGGT|CAGACC|CATCAGATTTCTG 2222
Db 66752 TAAGCA|AATTAG|CAATTC|CAGAGAA|CCTGAA|GGGGT|CAGACC|CATCAGATTTCTG 66811
QY 2223 TGCCAC|ATGCCAA|GTACTTCT|GAGCATG|ACTGGATG|AGCTGTCC|CATCTGAAATCATC 2282
Db 66812 TGCCAC|ATGCCAA|GTACTTCT|GAGCATG|ACTGGATG|AGCTGTCC|CATCTGAAATCATC 66871
QY 2283 CAGTCTT|GTTT|CAGAACTTT|CACA|CGGAC|AGGAG|CCAGGACTG|GAATG|CAGTCTCTCG 2342
Db 66872 CAGTCTT|GTTT|CAGAACTTT|CACA|CGGAC|AGGAG|CCAGGACTG|GAATG|CAGTCTCTCG 66931
QY 2343 TCAC|TGGC|CAGAGT|TGGCTT|GACCT|TGAG|ACCAGT|GGCCAA|CAAAAGGAGTGTCTTAG 2402
Db 66932 TCAC|TGGC|CAGAGT|TGGCTT|GACCT|TGAG|ACCAGT|GGCCAA|CAAAAGGAGTGTCTTAG 66990
QY 2403 TCTAC|CTCC|CAGGAAATCC|CAGTGTCTT|GTTCTTCT|GGGAAGTGAATCA|TTGGCGCAGCA 2462
Db 66991 TCTAC|CTCC|CAGGAAATCC|CAGTGTCTT|GTTCTTCT|GGGAAGTGAATCA|TTGGCGCAGCA 67050
QY 2463 CTCGTATTT|TCTCTCT|TCCAGG|GAAGATCTT|AGG|CAGTATTT|TGGGAAGCATG 2522
Db 67051 CTCGTATTT|TCTCTCT|TCCAGG|GAAGATCTT|AGG|CAGTATTT|TGGGAAGCATG 67110
QY 2523 GGCAT|GGAAG|GAC|CCGGT|GATG|CATG|CTGCTG|GTTCTG|AGCTCTCATG|GTAAG 2582

Db 67111 GGCAT|GGAAG|GAC|CCGGT|GAATG|CATAG|CTCCCT|CGCTGGT|TCTCAG|GCTCTCAT|GTAAG 67170
QY 2583 CTCT|CA|CAG|CA|CGG|AAAA|GATGGGG|C|CAGG|GAC|AGAT|CAGT|AGG|T|CAG|AGC|ATCTC 2642
Db 67171 CTCT|CA|CAG|CA|CGG|AAAA|GATGGGG|C|CAGG|GAC|AGAT|CAGT|AGG|T|CAG|AGC|ATCTC 67230
QY 2643 AGG|SAC|CGG|GGG|CAATAT|GCTCT|GAG|CAGG|GATTA|GAG|CTTGGG|CTCTCAT|ATG|GTG 2702
Db 67231 AGG|SAC|CGG|GGG|CAATAT|GCTCT|GAG|CAGG|GATTA|GAG|CTTGGG|CTCTCAT|ATG|GTG 67290
QY 2703 TTTCTGG|GCTCA|ACTG|CCAG|CTCCGT|CACTTACT|GTTG|TGTG|TGTG|TGTG|TGTG|TGTG|TGTG 2762
Db 67291 TTTCTGG|GCTCA|ACTG|CCAG|CTCCGT|CACTTACT|GTTG|TGTG|TGTG|TGTG|TGTG|TGTG 67350
QY 2763 CCAT|CTCT|CCAT|ATCTTT|CTTCT|CACTTTTAAAT|GGAATTAAT|GGG|GTAC|CCAC|CTCCCA 2822
Db 67351 CCAT|CTCT|CCAT|ATCTTT|CTTCT|CACTTTTAAAT|GGAATTAAT|GGG|GTAC|CCAC|CTCCCA 67410
QY 2823 GGT|CA|CAG|AGG|GCTTAC|AGAA|ACGATTTCTTGTGAATTTGGCTTGCAGTAAATTAATCAA 2882
Db 67411 GGT|CA|CAG|AGG|GCTTAC|AGAA|ACGATTTCTTGTGAATTTGGCTTGCAGTAAATTAATCAA 67470
QY 2883 TACT|TCC|AGCTATTTCTTATTTCCAT|TCCAT|TCCAT|TCCAT|TCCAT|TCCAT|TCCAT|TCCAT|TCCAT 2942
Db 67471 TACT|TCC|AGCTATTTCTTATTTCCAT|TCCAT|TCCAT|TCCAT|TCCAT|TCCAT|TCCAT|TCCAT 67530
QY 2943 CATG|TCAG|TGTTCCTG|ACG|GTTTCC|AAAG|AGATTTCC|AAAG|AGATTTCC|AAAG|AGATTTCC|AAAG 3002
Db 67531 CATG|TCAG|TGTTCCTG|ACG|GTTTCC|AAAG|AGATTTCC|AAAG|AGATTTCC|AAAG|AGATTTCC|AAAG 67590
QY 3003 TGAAG|AATCTCC|AAAC|ATCC|CGC|ACGATCC|TGTG|GCG|CAGG|CACA|CACA|CAGC|CAGC|CTGTG 3122
Db 67591 TGAAG|AATCTCC|AAAC|ATCC|CGC|ACGATCC|TGTG|GCG|CAGG|CACA|CACA|CAGC|CAGC|CTGTG 67650
QY 3063 CCG|CC|CGG|GCTCTG|AA|CAGG|ATG|GCTG|CG|CAGG|CACA|CACA|CAGC|CAGC|CTGTG 3122
Db 67651 CCG|CC|CGG|GCTCTCTG|AA|CAGG|ATG|GCTG|CG|CAGG|CACA|CACA|CAGC|CAGC|CTGTG 67710
QY 3123 TGTG|CGG|CGG|GAGTCCG|GTG|CGG|TCC|CGG|TGCAG|CGG|TGCAG|CGG|TGCAG|CGG|TGCAG 3182
Db 67711 TGTG|CGG|CGG|GAGTCCG|GTG|CGG|TCC|CGG|TGCAG|CGG|TGCAG|CGG|TGCAG|CGG|TGCAG 67770
QY 3183 AGC|ATTTGTT|CGC|AGG|CGTAC|CGG|AGC|CCCG|CGCTG|CGC|CGG|GAGG|GAGG|GAGG|GAGG|GAGG 3242
Db 67771 AGC|ATTTGTT|CGC|AGG|CGTAC|CGG|AGT|CCCC|CGCTG|CGC|CGG|GAGG|GAGG|GAGG|GAGG|GAGG 67830
QY 3243 CCG|GT-|CCCC|AAG|CTC|AGAT|CTG|GGT|GGT|GCC|AC|GCTCT|CCCT|GCC|AC|GCTCTG|GCC|AC|GCTCTG 3301
Db 67831 CCG|GT-|CCCC|AAG|CTC|AGAT|CTC|GGT|GGT|GCC|AC|GCTCT|CCCT|GCC|AC|GCTCTG|GCC|AC|GCTCTG 67890
QY 3302 GGG|GAC|GGG|AAG|AGCGG|ACCG|GAGATTTAGT|GTTGG|GGG|CCCC|CGG|AGG|GTTCA|C|CACTG 3361
Db 67891 GGG|GAC-|GGG|AAG|CGG|ACCG|GAGATTTAGT|GTTGG|GGG|CCCC|CGG|AGG|GTTCA|C|CACTG 67949
QY 3362 TTTCT|GAG|AAA|CTTCC|CACTG|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA 3421
Db 67950 TTTCT|GAG|AAA|CTTCC|CACTG|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA 68009
QY 3422 CTGGG|CTAG|GCTCGG|CGCC|CAAG|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA 3481
Db 68010 CTGGG|CTAG|GCTCGG|CGCC|CAAG|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA|CCCA 68069
QY 3482 CTCCC|CAG|CGG|GATCCCG|GGC|AGG|CCCG|CGG|GGTGG|AGAG|AGCTGAG|AGGA 3541
Db 68070 CTCCC|CAG|CGG|GATCCCG|GGC|AGG|CCCG|CGG|GGTGG|AGAG|AGCTGAG|AGGA 68129
QY 3542 GAAA|CAG|AGG|GGG|GGG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG 3601
Db 68130 GAAA|CAG|AGG|GGG|GGG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG 68189
QY 3602 CCAAT|GCG|AA|CGG|CAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG 3661
Db 68190 CCAAT|GCG|AA|CGG|CAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG|GAG 68249


```

QY 3662 TTTACAGAGCCCTCATTAGCATTTCCAGAGAGGAGGGGAGGGGAGAGGCGGGTGT 3721
Db 68250 TTTACAGAGCCCTCATTAGCATTTCCAGAGAGGAGGGGAGGGGAGAGGCGGGTGT 68309
QY 3722 GTGGTGTGGTGTGGGAGCATCCCGGCGCCCTGTGTGGGTGCGCGGAGCCCTCGGCCCT 3781
Db 68310 GTGGTGTGGTGTGGGAGCATCCCGGCGCCCTGTGTGGGTGCGCGGAGCCCTCGGCCCT 68369
QY 3782 GTGTCTCTCTCCCTCCCGCCCTTACTCTCCAGCGGGAGCCGCGCGGAGTCACTCCCT 3841
Db 68370 GTGTCTCTCTCCCTCCCGCCCTTACTCTCCAGCGGGAGCCGCGCGGAGTCACTCCCT 68429
QY 3842 CGGACTTTGCGCCCTGTGTGGCAGCGGATAAAGGGGGCTGAGGAATACCGGACACGGTC 3901
Db 68430 CGGACTTTGCGCCCTGTGTGGCAGCGGATAAAGGGGGCTGAGGAATACCGGACACGGTC 68489
QY 3902 ACCCGTTGC 3910
Db 68490 ACCCGTTGC 68498

RESULT 5
AL391723/c
LOCUS AL391723 185644 bp DNA linear HTG 11-SEP-2001
DEFINITION Homo sapiens chromosome 10 clone RP11-38B21, WORKING DRAFT
ACCESSION AL391723
VERSION AL391723.7 GI:10186824
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Burton, J.
Direct Submission
Submitted (10-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk
On Sep 19, 2000 this sequence version replaced gi:10178830.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA38B21
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: pBS1972; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 175002 bases at least Q40
Consensus quality: 179431 bases at least Q30
Consensus quality: 182156 bases at least Q20
Insert size: 184344; sum-of-contigs
Insert size: 140788; 6.0% error; agarose-fp
Quality coverage: 4.15x in Q20 bases; sum-of-contigs Quality
coverage: 6.51x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 59110: contig of 59110 bp in length
* 59111 59210: gap of 100 bp
* 59211 61285: contig of 2075 bp in length
* 61286 61386 64379: gap of 100 bp
* 61386 64379: contig of 2994 bp in length

```

```

* 64380 64479: gap of 100 bp
* 64480 66747: contig of 2268 bp in length
* 66747 66847: gap of 100 bp
* 66847 66848 114355: contig of 47508 bp in length
* 114355 114356: gap of 100 bp
* 114356 125406: contig of 10951 bp in length
* 125406 125407: gap of 100 bp
* 125407 125507: contig of 2167 bp in length
* 125507 127773: gap of 100 bp
* 127773 130111: contig of 2338 bp in length
* 130111 130212: gap of 100 bp
* 130212 132354: contig of 2143 bp in length
* 132354 132355: gap of 100 bp
* 132355 136537: contig of 4083 bp in length
* 136537 136538 139127: contig of 2490 bp in length
* 139127 139128 139228: gap of 100 bp
* 139228 181355: contig of 42128 bp in length
* 181355 181356: gap of 100 bp
* 181356 183524: contig of 2069 bp in length
* 183524 183525: gap of 100 bp
* 183525 185644: contig of 2020 bp in length.
* 185644 185645: gap of 100 bp
* 185645 185646: contig of 2020 bp in length.

FEATURES
Location/Qualifiers
1..185644
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-38B21"
/clone_lib="RPC1-11.1"
1..59110
/note="assembly_fragment:02307"
clone_end:17
vector_side:left
59211..61285
/note="assembly_fragment:00197"
fragment_chain:1
61386..64379
/note="assembly_fragment:01916"
fragment_chain:1
64480..66747
/note="assembly_fragment:00079"
fragment_chain:1
66848..114355
/note="assembly_fragment:00585"
fragment_chain:2
114356..125406
/note="assembly_fragment:02383"
fragment_chain:2
125407..127773
/note="assembly_fragment:00150"
fragment_chain:1
127774..130111
/note="assembly_fragment:00464"
fragment_chain:1
130112..132354
/note="assembly_fragment:00629"
fragment_chain:1
132355..136537
/note="assembly_fragment:00814"
fragment_chain:1
136538..139127
/note="assembly_fragment:00861"
fragment_chain:1
139128..181355
/note="assembly_fragment:02212"
fragment_chain:1
181356..183524
/note="assembly_fragment:02327"
fragment_chain:1
183525..185644
/note="assembly_fragment:02419"
fragment_chain:1

```

ORIGIN

```

Query Match 45.5%; Score 1779.2; DB 2; Length 185644;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1823; Conservative 0; Mismatches 118; Indels 3; Gaps 3;
QY 1968 GTCTCTACTAAATAACAAATAATTAGTAAAGTGTGGTGGCGCATGCTGTATCCAGC 2027
|||||

```

[illegible]

QY	3108	ACCAGCCAGCCTGTGTGTGTCGGCCGCGAGTCCCGTGCGGTCCCGGGTGACAGCGCGTGCC	3167
Db	182505	ACCAGCCAGCCTGTGTGTGTCGGCCGCGAGTCCCGTGCGGTCCCGGGTGAGCAGCGGTGCG	182446
QY	3168	TGGTGGGCGGGGAGAGCCATTGTTCGAGGGGTACCCGAGCCCGCCCGCGCTGCGCCCGGA	3227
Db	182445	TGGTGGGCGGGGAGAGCCATTGTTCGAGGGGTACCCGAGTCCCGCGCGCTGCGCCCGGA	182386
QY	3228	GGGAGCGGGGCTTCCCGGT-CCCCAGCTCCAGTCTCTGGGTGGCTGCCACAGTCTCC	3286
Db	182385	GGGAGCGGGGCTTCCCGGTCCCCAAGCTCCAGTCTCTGGGTGGCTGCCACAGTCTCC	182326
QY	3287	CTGCCACGCGCTGTGGGGGACGGGAAGACGGGACGGAGATGTAGTGTGGCGGCCCGCC	3346
Db	182325	CTGCCACGCGCT- GGGGGACGGGAAGACGGGACGGAGATGTAGTGTGGCGGCCCGCC	182267
QY	3347	GAGGTTTACCACTGTCTTCTGAGAACTTCCCGAGTCCCAACCCAGCTTCTCCGT	3406
Db	182266	GAGGTTTACCACTGTCTTCTGAGAACTTCCCGAGTCCCAACCCAGCTTCTCCGT	182207
QY	3407	GCCCAGGGCGGTCCTGGGCTTAGCTCCGCGCCCGCCAGCCCCAAACGGGTCCCCAGCCC	3466
Db	182206	GCCCAGGGCGGTCCTGGGCTAGCTCCGCGCCCGCCAGCCCCAAACGGGTCCCCAGCCC	182147
QY	3467	CTTCCAGAGAGAAGCTCCCGACGCGGATGCCGGGACAGGCCCGAGCGGGTGGAA	3526
Db	182146	CTTCCAGAGAGAAGCTCCCGACGCGGATGCCGGGACAGGCCCGAGCGGGTGGAA	182087
QY	3527	AGAACTCAGAAAGAGAAACAGAGGGAGGGGGAGCGAGAGCTGCGCGGACAGGGCCAG	182087
Db	182086	AGAACTCAGAAAGAGAAACAGAGGGAGGGGGAGCGAGAGCTGCGCGGACAGGGCCAG	182027
QY	3587	GCAGATTGCGCCGAGCCAAATGGCAACGGCAGGACGAGGTGCACCAATTCCTTCGCCC	3646
Db	182026	GCAGATTGCGCCGAGCCAAATGGCAACGGCAGGACGAGGTGCACCAATTCCTTCGCCC	181967
QY	3647	AATCAGCAGCGGATTTACAGAGCCTCATAGCATTTCCACAGAGGACGGGACAGGG	3706
Db	181966	AATCAGCAGCGGATTTACAGAACTCATATAGCATTTCCACAGAGGACGGGACAGGG	181907
QY	3707	CAGAGCGGGTGTGTGTGTGTCGGTGTCCGACAGCATCCCGCGGCCCTCTCGCGTCC	3766
Db	181906	CAGAGCGGGTGTGTGTGTGTCGGTGTCCGACAGCATCCCGCGGCCCTCTCGCGTCC	181847
QY	3767	CGCAGCCTCGGCTCTGTCTCTCCCTCCCGCCTTACCTTCACGCGGGACCGCCCG	3826
Db	181846	CGCAGCCTCGGCTCTGTCTCTCCCTCCCGCCTTACCTTCACGCGGGACCGCCCG	181787
QY	3827	CGCCAGTCAACTCTCGCACTTTTGCCCTGCTTTGGCAGCGGATAAAAGGGGGCTGAGGA	3886
Db	181786	CGCCAGTCAACTCTCGCACTTTTGCCCTGCTTTGGCAGCGGATAAAAGGGGGCTGAGGA	181727
QY	3887	ATACCGGACAGGTCAACCGTTGC	3910
Db	181726	ATACCGGACAGGTCAACCGTTGC	181703

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Cercopithecinae; Papio.
1 (bases 1 to 161827)
Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
Peng, Z., Malinov, I. and Rubin, E.M.
Direct Submission
Unpublished
2 (bases 1 to 161827)
Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
Peng, Z., Malinov, I. and Rubin, E.M.
Direct Submission
Submitted (08-FEB-2003) Genome Sciences, Lawrence Berkeley National
Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA

Sequence Produced by Berkeley PGA
Web site: <http://pga.lbl.gov>
Center Code: EGABERK
Center Project Name: B051
Bac Clone Name: RP41-194B17

This sequence has been compared to sequences of other species
using Vista (<http://www.gsdlbl.gov/VISTA>). The results can be
viewed at:
http://pga.lbl.gov/cgi-bin/search_cvgcd?type=nkvalue=SCD

The order-orientation of the draft sequence was accomplished by
using:
Avid (<http://baboon.math.berkeley.edu/mavid/>),
Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

Summary Statistics:
Sequencing vector: Plasmid; pUC18
Chemistry: Dye-terminator Big Dye
Assembly program: Phrap version 0.990329.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*
1 161827: contig of 161827 bp in length.
Location/Qualifiers
1. 161827
/organism="Papio hamadryas"
/mol_type="genomic DNA"
/db_xref="taxon:9557"
/clone="RP41-194B17"

FEATURES
source

ORIGIN

Query Match 43.8%; Score 1711.6; DB 2; Length 161827;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 1896; Conservative 0; Mismatches 134; Indels 24; Gaps 8;

QY 1858 GCCAGGCGTGGTCTATCCGCTGTAATCCAGCACTTTGGAGGGCCAAAGGCGATGGGAT 1917
DB 139308 GCTGGGCACGGTGGCTATCCGCTGTAATCCAGCACTTTGGAGGGCTAAGAGGGCGGAT 139367

QY 1918 CACCTGAGGTCAGAGGTTCAAAACCGAGCTGGCCAAATGGTGAACCCCGCTCTACTA 1977
DB 139368 CA--TGAGGTCAGGAATGCAAGACCGAGCTGACCAATAGGGTGAACCCCGCTGACTA 139425

QY 1978 AAAATCAAAAAATAGCTAAGTGCTGGTGGCGATGCCGTGTAATCCAGCTATTGGGAG 2037
DB 139426 AAATACAAAAATAGCTAAGTGCTGGTGGCGATGCCGTGTAATCCAGCTATTGGGAG 139485

QY 2038 GGTGAGGCGGAGAAATTTCTTGAACCGGAGGAGCGAGGTTGCAAGTAAGCGAGATCA 2097
DB 139486 GCTGAGGCGAGGAAATTTCTTGAACCGGAGGAGCTGAGGTTGCAAGTAAGCGAGATCA 139545

QY 2098 CCACCTGCACTCCAGCCCTGGGGGAGAGCGAGACTTCTCTCAAAAAACAAAAACAA 2157

DB 139546 CCACCTGCACTCCAGCGCTGGGGACAGCGAGCTCCCTCTCAAAAAACAAAAACAA 139605
QY 2158 AGAATTAAGCAAAATAGACATTCGAGAGAGACCTGAAGGGGTGAGACCAGTACAGAT 2217
DB 139606 AATATTCAGCAAAATAGACATTCGAGAGAGACCTGAAGGGGTGAGACCAGTACAGAT 139665
QY 2218 TTCTGTGCACATGCCAAATAGACATTCGAGAGATGATGATGAGCTGTGCCACATCTGAAA 2277
DB 139666 TTCTGTGCACATGCCAAATAGACATTCGAGAGATGATGATGAGCTGTGCCACATCTGAAA 139725
QY 2278 TCATCCACTCTGTTCAGAACTTTTCACCCGAGAGGAGCCAGGACTGGAATGCAGTCT 2337
DB 139726 TCATCCAAATCTGTTCAGAACTTTTCACCCGAGAGGAGCCAGGACTGTATGAGCT 139785
QY 2338 CCTGTCACTGGCCAGAGAGTTGGCTTACCTTCAGACCTGAGACCTGAGCCCAACAAAGAGCTG 2397
DB 139786 CCTGTCACTGGCCAGAGAGTTGGCTTGAACCTGAGACCTGAGCCCAACAAAGAGCTG 139844
QY 2398 CTTAGTCTACCTCCAGGAAATCCAGGTGCTTCTTCTTGGGAAGTGAATCAATTGAGCT 139892
DB 139845 ATTAGTCTTCTTCCAGGAAATCCAGG-----TGGGAGCGAATCAATTAGCG 2457
QY 2458 CAGCACTCCGTATTTTCTCTTCCAGGGGAAGATCTTAGGGCAGTATTTGGGAAAG 2517
DB 139893 CAGCACTCCGTATTTTCTCTTCCAGGGAATGATCTTAGGGCAGAAATTTGGGGAG 139952
QY 2518 ACATGGGCTGGAAGAGACACCGGGTGAATGCATAGCTGCTGCTGAGCTCTCATGG 2577
DB 139953 ACATGGGCTGGAAGAGAGAGGGGTGAATGCATAGCTGCTGCTGAGCTCTCATGG 140012
QY 2578 TAAGGCTCTACAGACACCGGAAAGATGGGGGCAACAGGACAGATAGTAGGGTCAAGC 2637
DB 140013 TAAGACTCTACAGACACCGGAAAGATGGGGGCAACAGGACAGATAGTAGGGTCAAGC 140072
QY 2638 ATCTCGGACCGGAGGCAATATGCTCTCAGCAGAGGATTAAGAGCTTGGGCTCTCATAT 2697
DB 140073 ATCTCGGACCGGAGGACAGTATGGTCTGTAACAGGAAATTAAGAGCTTGGGCTCTCATAT 140132
QY 2698 GGTGTTTCTGGGCTCAACTGCCAGCTCCGTCATCTACTGCTGCTGACATGGGCAAG 2757
DB 140133 GGCATTTCTGGGCTCAATTCGCACTCCATCACTTACTGCTGCTGATCATGGGCAAG 140192
QY 2758 TTATTCATCTCTCATATCTCTTCTCTCACTTTTAAATGGAATTAATGGGTACCCACC 2817
DB 140193 TTATTCATCTCTCTATATCTTTCCACACTTATAAATGGAATTAATGGGTACCCACC 140252
QY 2818 TCCAGGCTCACAGAGAGGTTACAGAAACGATTTCTTGAATTTGGCTTGCAGTAAATA 2877
DB 140253 TCCAGGCTCACAGTGGAGATTACAGAAACGATTTCTTGAATTTGGCTTGCAGTAAATA 140312
QY 2878 TTCAATPACCTGCCAGCTATTCTTATTTCCACATCCAGCCCTTTGGCTGCTGCTGGGTGA 2937
DB 140313 TTCAATPACCTGCCAGCTATTCTTATTTCCACATCCAGCCCTTTGGCTGCTGCTGGGTGA 140372
QY 2938 AAACATGTCAGTGTCTTCTGACGGTTTCCAAAGAAAGATTCCAAATTAACCTTGC 2997
DB 140373 AAGACATGTCAGTGTCTTCTGACGTGTTTCCACA---AAGATTCCAAATTAACCTTGC 140429
QY 2998 CAGTCTGAAGAAATCTCCAAACATCCCGACGATCTCGAGCGCGGGCTTTGGGATGG 3057
DB 140430 CAGTCTGAAGAAATCTCCAAACATCCCGACGATCTCGAGCGCGGGCTTTGGGATGG 140489
QY 3058 GACTGCCCGCGGCTCTTGAACAGGATGCGTGCAGGAGGACACACACACAGCCAGC 3117
DB 140490 GACTGTTTGGCGGGGCTGAAACAGATTCGTCGCGGTAGG--CACACACACAGCCAG 140547
QY 3118 CTGTGTGCGGCGGAGTCCGGTCCGGTCCGGTGTAGAGCGCGTGGTGGGCGG 3177
DB 140548 C--TCGCTGCGGCGGAGTCCGGTCCGGTGTAGAGCGCGTGGTGGGCGG 140605
QY 3178 GGCAGAGCAATTTGTCGAGGCGTACCGAGCCCGCGCTGCCCGGAGGGAGCGGG 3237

Db 140606 GGCAGAGCCATTGTTCCCGAGCTGACTGAGCCCCCAGCGCTCGCCCCGGGAGGAGCGGG 140665
QY 3238 GCTTCCCGCGTCCCC-AAGTCCAGATCCCTGGGGTGGCTGCCAGTCTCCCTGCCACGCG 3296
Db 140666 GCTTCCCGCGTCCCGAGCTCCCGATCCCGGGTGGCGGCTCGTCTCCCTGCCACGCG 140725
QY 3297 CTTGGGGGACGGAGACGGGACGAGAGTGTAGTGTGGGCCCCCCCCCGAGGTTCCAC 3356
Db 140726 CTTGGGGGACGGAGACGGGACGAGATGTGTAGTAGTGGGCTCTCCCGAGGCTTCCC 140785
QY 3357 CACTGTTTCTCAGAAATTTCCCGAGTGCACCCACCCCTCTCGTGTGCGCGAGGGC 3416
Db 140786 CACTGTTTCTCAGAAATTTCCCGAGTGCACCCACCCCTCTCGTGTGCGCGAGGGC 140845
QY 3417 CGGTCTGGGCTAGGCTCGCGCCCGACCGCCCAACCGGGTCCCGAGCCCTTCCAGAGA 3476
Db 140846 CGGTCTGGGCTAGACTCGCGCGCCCGACCGCCCAACCGGGTCCCGAGTCCCTTCCAGAGA 140905
QY 3477 GAAAGTCTCCGAGCGCGGATCGCGGCGAGAGCCCGCGCGGTGGAAGAGAGTGGAG 3536
Db 140906 GAAAGTCTCTAACCGCGGGTTCGCGGCGAGAGCCCGCGCGGTGGAAGAGAGTGGAG 140965
QY 3537 AAGGAGAAACAGAGGGGAGGAGGAGAGTGGCGGCGAGAGGAAACAGCAGATTGCG 3596
Db 140966 AAGGAGAAAGGAGGAGGAGGAGGAGAGTGGCGGCGAGAGGAAACAGCAGATTGCG 141025
QY 3597 CCGAGCAATGGCAACGGGAGAGAGGAGTGGCAACCAATTCCTTCGGCAATGACGAGC 3656
Db 141026 CCGAGCAATGGCAACTGCGAGAGAGTGGCAACCAATTCCTTCGGCAATGACGAGC 141085
QY 3657 CGAGTTTACAGAGCTCATTTAGCATTTTCCCGAGGCGAGGGGCGAGGCGCGG 3716
Db 141086 CGAGTTTACAGAGCTCATTTAGCATTTTCCCGAGGCGAGGGGCGAGGCGCGG 141145
QY 3717 GTGGTGTGTGTGTCGGTGTGCGAGCATCCCGCGGCTGTGTCGGTTCGCGGAGCCTC 3776
Db 141146 GTGGTGTGTGTGTCGGTGTGCGAGCAT-CCCGCGGCTGTGTCGGTTCGCGAGCCTT 141204
QY 3777 GGCCTGTCTCTCTCCCGTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 3836
Db 141205 GGCCTGTCT 141264
QY 3837 CTCTTCGCACTTTGCCCTGTGTCAGCGGATTAAGGGGCTGAGGAATACCGGACA 3896
Db 141265 CTCCGCGACTTTGCCCTGTGTCAGCGGATTAAGGGGCTGAGGAATACCGGACA 141324
QY 3897 CGGTCAACCGTTGC 3910
Db 141325 CGGTCAACCAATGC 141338

RESULT 8
LOCUS AC026883
DEFINITION Homo sapiens chromosome 3 clone RP11-799L19 map 3p, WORKING DRAFT
SEQUENCE, 41 unordered pieces.
ACCESSION AC026883.3 GI:8101285
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 179879)
AUTHORS Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Guan, Q., Gu, X., Guo, D., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, S., Sun, M., Sun, W., Sun, Y., Tao, R., Wang, H., Wang, J., Wang, L., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y.,

TITLE
JOURNAL
REFERENCE
AUTHORS
Zhang, Z., Zhu, B., Yu, J. and Yang, H.
Chromosome 3p genomic sequence
Unpublished
2 (bases 1 to 179879)
Li, W., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y., Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y., Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D., Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J. and Yang, H.
Direct Submission
Submitted (25-MAR-2000) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China
On May 29, 2000 this sequence version replaced gi:7417472.
-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hgsc.igtp.ac.cn
http://www.genomics.org.cn
Contact:hgsc@igtp.ac.cn
-----Summary Statistics
Center project name:if project
Center clone name: RP11-799L19
-----Project Information
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator: ET 55% of reads
Chemistry: Dye-terminator: Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 149369 bases at least Q40
Consensus quality: 160276 bases at least Q30
Consensus quality: 167926 bases at least Q20
Insert size: 148122; sum-of-contigs
Quality coverage: 3.56x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 41 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1
1497: contig of 1496 bp in length
1596: gap of unknown length
1597: contig of 1353 bp in length
2949: gap of unknown length
3049: gap of unknown length
3050: contig of 1360 bp in length
4409: gap of unknown length
4510: contig of 1628 bp in length
6138: gap of unknown length
6238: contig of 1047 bp in length
7285: gap of unknown length
7385: contig of 1629 bp in length
9014: gap of unknown length
9114: contig of 1331 bp in length
10444: gap of unknown length
10545: contig of 1609 bp in length
12154: gap of unknown length
12254: contig of 2133 bp in length
14386: gap of unknown length
14487: contig of 2504 bp in length
16990: gap of unknown length
17090: contig of 1932 bp in length
17091: gap of unknown length
19023: contig of 1375 bp in length
19123: gap of unknown length
20498: contig of 1612 bp in length
20597: gap of unknown length
22309: contig of 2368 bp in length
22310: gap of unknown length
24677: gap of unknown length
24778: contig of 2878 bp in length
27655: gap of unknown length
27755: gap of unknown length

```
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
```

QY	3	GAATTCATCCATTTAAATCATACAATTTAAATGGCTTTTAGTATATTTACAGGTTGTGCAT	62
Db	10955	GAATTCATCCATTTAAATCATACAATTTAAATGGCTTTTAGTATATTTACAGGTTGTGCAT	110014
QY	63	CCATCACAAATCCATTTTAGACAGTTTATTACTCCAAAAATAAACCTGCAATTCCTTAG	122
Db	110015	CCATCACAAATCCATTTTAGACAGTTTATTACTCCAAAAATAAACCTGCAATTCCTTAG	110074
QY	123	CCATCACCCCCCAACATCCTCCATCCTCCTTCCAAAGCCCTGGGCAACACCAAACTCACTT	182
Db	110075	CCATCACCCCCCAACATCCTCCATCCTCCTTCCAAAGCCCTGGGCAACACCAAACTCACTT	110134
QY	183	TCTGTCTCTATAAATTTGCCAAATTCCTGGACATTTCATATAAATGGAAGCAAAACAACATGT	242
Db	110135	TCTGTCTCTATAAATTTGCCAAATTCCTGGACATTTCATATAAATGGAAGCAAAACAACATGT	110194
QY	243	GAGACTTTGTGACTGGCTGCTTTTACATTAGCATTTCTATTTTAAAGGCTCATTTATGTTACA	302
Db	110195	GAGACTTTGTGACTGGCTGCTTTTACATTAGCATTTCTATTTTAAAGGCTCATTTATGTTACA	110254
QY	303	GTACTTAGCAGTACTTCATTTCTTTTTTATTCTCAAATGGTATTCCACTGTGTGGGTATCC	362
Db	110255	GTACTTAGCAGTACTTCATTTCTTTTTTATTCTCAAATGGTATTCCACTGTGTGGGTATCC	110314
QY	363	CATATCATATTATTAGACAGGTTCTCACTGTGCACCCAGGCTGGAGTGCAGTGGCAC	422
Db	110315	CATATCATATTATTAGACAGGTTCTCACTGTGCACCCAGGCTGGAGTGCAGTGGCAC	110374
QY	423	AATCAUAGTCTACGTAAACCTCAAACTCCCTGGGCTCAAGTGATTCCTACTACCTCAGGCTC	482
Db	110375	AATCAUAGTCTACGTAAACCTCAAACTCCCTGGGCTCAAGTGATTCCTACTACCTCAGGCTC	110434
QY	483	CAGAGTAGCTAGGACTACAGGCACACAGCCATACCTGGCTAAATTTTTTTTTTTTAAATTT	542


```
QY 2244 GAGGATGCTGGATGAGTGTCCATCTGAAATCATCCAGTCTTGTTCAGAACTTTCA 2303
Db 120191 GAGTCAATTTCTAGATGAGTGTCCATCTGAAACATCTAGTCCCATCTGAACTTTCA 120250
QY 2304 CACCGGACAGGAGCAGGACCTGGAATGAGTCTCTCGTCACTGCGCAGAGAGTTGGCC 2363
Db 120251 CACTGGAAGAAAGCAGGACAGGAAATAGACATCTGCTCACTGGCCAGAAAGTTGTC 120310
QY 2364 TTGACCTTGAGACCACTGCGCAACAAAGAGAGTCTTAGTCTAATCTCCAGGAAT-CCC 2422
Db 120311 -TGACCCAGAGACAGCGCCCAACAGGAGAGTCTTAGTCTCTCTGCGCAGGCAATCCCC 120369
QY 2423 AGGTGCTTGTCTCTCTGGAAGTGAATCATTTGGCGACATCTCGGTATTTTCTCTCTTC 2482
Db 120370 AAGTGTCTCTCTCTGGAAGCAGATGGGTGGCACAACTTCTTATTTTGTCTCTTC 120429
QY 2483 CCAGGGAGAGATCTTAGGCGAGTATTTGGGAAGACATGCGGATGGAAGACACCGGT 2542
Db 120430 CCAGGAAGAGATCTTGGGCGAGTATTTGGG--GAGACAGCAGGGAAGAGACTGGT 120486
QY 2543 GAATGATAGCTGCTGCTGTTCTGAGCTCTCATGTAGGCTCTTACAGACCGAAAG 2602
Db 120487 GAATACAGAGCTGCTGCTGTTCTGTTACGTGAGGCTCTCCACAGGCGCCAGAAAT 120546
QY 2603 ATGGGGCAGAGGACAGATCAGTAGGCTCAGAGCATCTCAGGACCGCAGGCAATATGG 2662
Db 120547 ATAGGGGCAAGAG-----AGACAGTCAAGACATCTCAGGACTAGGGCC----- 120591
QY 2663 TCCTGACAGGATTAAGACTTGGGCTCTCATATGTTGTTTCGCGCTCAACTGCCAGC 2722
Db 120592 TTCTGAGTGGGATTAAGAGCGCGCGGCTTGCAATTACATTTCTGGG--CAATTCCTAGC 120649
QY 2723 TCCTGACTTACTGTTGCTGTGACCATGAGCAATGAGTATTCATCTCTCATATCT--CT 2780
Db 120650 TTCACCACTTACGCTGTGCTGTGACCATGAGTATTCATCTCTCATATCT--CT 2780
QY 2781 TTCTCTACTTTTAAATGGAATAGTGGGTACCCACTCCAGGGTCACAGAGAGGCTTA 2840
Db 120710 TTCTCTACTTTTAAATGGAATAGTGGGTACCCACTTCATAGGTACAGTGAAGAT-- 120768
QY 2841 CAGAAACGATTTCTTGATTTGGCTTGAGTGAATTAATCAATCACTGCCAGCTATCTT 2900
Db 120769 ----AATGAATCTGTGAATGGCTGCGCAGTAAACAAATTCATATCTGTAACTATCTT 120824
QY 2901 ATTCACATCCAGCCCTTTGCGCTGCTGCTGAGTGAACACATCTCAAGTGTTCCTGA 2960
Db 120825 -TTCCACCTAGCAGTCTTTGCTGATCTGGGTGAAGCAGTCTGTGCTTCTGCT 120883
QY 2961 CGGTTTCCAGAAAGATTCAAAATTAACCTG--CCAGTCTGAAGATCTCCAAAAC 3019
Db 120884 TTTTTCATTAAGA--GTTCCAAAATTAACCTGACAAATCTGAAGATCTCCAAAAC 120941
QY 3020 ATCCCGCAGCATCTTGGAGGCGCGGCTTGGGATGAGCTGCCCGCGGCTGCTGA 3079
Db 120942 GTCCCGCAGCGTTCTAGAGGCGTGGGCGACGGGTCGCTGCGCCACTGGTGGCGG 121001
QY 3080 CAGGA-----TGCGTGGCGCAGGACACAC 3105
Db 121002 GTGACGCTCCCTGCGAGCGGTACCCAGTGGGGCACAGGTGGCGCGCAGGACACAC 121061
QY 3106 ACACAGCCAGCTGTGTGTG-----CGGCGGAGTCCGGTGGCGTCCCGGTGAGCAG 3159
Db 121062 GTACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 121120
QY 3160 CGCGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 121180
Db 121121 CGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 121180
QY 3220 GCCCGGAGGAGCGCGGCTTCCCGGCTCCCAAGCTCCAGATCTCTGGGTGGCTGCCA 3279
Db 121181 GCGGACAGAGGCGCGGCTTCCCGGCTTCCCGGCTTCCCGGCTTCCCGGCTTCCCGGCTTCC 121239
QY 3280 CGTCTCTCTGCGACGCGCTGTGGGGGAGCGGGAAGACGGGACGGAGATTTAGTGTGGG 3339
```

```
Db 121240 GTCTGCTCTCCCGCGCGCTTGGGGGAGCGGAATGCGGGAAGAGTCTGTGAGGATGGC 121299
QY 3340 GCCCGCGGAGGTTTCCACACTGTTTCTTGAGAAATTTCCCGAGTGGCCACCCACCGGTTCC 3399
Db 121300 G-CAACGCTGCTGCCACTGTTTCTGTGAAATTTGCCAGTGGCCACCGCA----TTC 121354
QY 3400 TCGTGTGCCGAGGCGGCTCTGCTGCTAGGCTCCGCGCCCGCCAGCCCAACCGGGTCC 3459
Db 121355 GCGCTGTGCCCGGCGC---ACCTGAGCTTGGAGCCACCCCGAGTTCCAGAACGATCC 121411
QY 3460 CCAGCCCTTCCAGAGAAAGTCTCCGACGCGGATGCGGGGAGAGGCGCCAGCGGCGG 3519
Db 121412 GCTCCCTCTCACACAGGAGAGTCTCCGA-----GGCAGGCGCCAGCGCGG 121458
QY 3520 GTGGAAGAGAGCTGAGAAGGAGAAACAGAGGGGAGGAGGAGGAGCTGGCGGAGAG 3579
Db 121459 GGGAGAGAGAGCGG---AGGAGAAAGGAGGAGGAGGAGGAGCTCGCGGAGAG 121515
QY 3580 GGGAAACAGCAGATTGCGCCGAGCCATGCGCAACGCGAGCAGAGTGGCACCAATTCCTCC 3639
Db 121516 GGGAAACAGCAGATTGCGCCGAGCCATGCGCAACGCGAGCAGAGTGGCACCAATTCCTCC 121575
QY 3640 TTGCGCCATGACGAGCGCGGAGTTTACAGAGCTCATTAGCATTTTCCCGAGGCGAGGG 3699
Db 121576 TTGCGCCATGACGCGCGCAGAGTTTACAGAGCTCATTAGCATTTTCCCGAGGCGAGGG 121635
QY 3700 GCAGGGCGCAGAGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3759
Db 121636 GCAGGGCGCAGCGC-----GGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 121690
QY 3760 CGGTCCCGCGAGCTCGGC-CTGTGTCTCTCTCC-CTCCCGCGCTTACCTCCAGCGG 3817
Db 121691 TGGTAGCTGCAGCGCTGGGCTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 121750
QY 3818 GACCGCGCGCGCAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3877
Db 121751 GACTGCTTGGCGCAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 121810
QY 3878 GTTGAGAAATACCGGACACGCTCACCGCTTGC 3910
Db 121811 TCTGAGAAATACCGGACACGCTCACCGCTTGC 121843
AY241932 17088 bp DNA linear MAM 02-APR-2003
Bos taurus stearoyl-CoA desaturase variant A (SCD) gene, complete cds.
ACCESSION AY241932
VERSION AY241932.1 GI:29469126
KEYWORDS Bos taurus (cow)
SOURCE Bos taurus
ORGANISM Bos taurus
REFERENCE 1. (bases 1 to 17088)
AUTHORS Medrano, J.F., Islas-Trejo, A.D. and Johnson, A.M.
TITLE Genomic sequence of the bovine stearoyl-CoA desaturase
JOURNAL Unpublished
REFERENCE 2. (bases 1 to 17088)
AUTHORS Medrano, J.F., Islas-Trejo, A.D. and Johnson, A.M.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2003) Animal Science, University of California, One Shields Avenue, Davis, CA 95616-8521, USA
FEATURES
source
1. .17088
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/chromosome="26"
1. .17088
gene
```



```
promoter /gene="SCD"
1..1880
mRNA /gene="SCD"
Join(1881..2051,2568..2850,6720..6850,8314..8519,
10099..10331,13019..17088)
CDS /gene="SCD"
/product="stearyl-CoA desaturase variant A"
Join(2051..2051,2568..2850,6720..6850,8314..8519,
10099..10331,13019..13218)
/gene="SCD"
/EC number="1.14.99.5"
/note="delta-9 desaturase"
/codon_start=1
/product="stearyl-CoA desaturase variant A"
/protein_id="AA063569.1"
/db_xref="GI:29469127"
/translation="MPAHLLOEISSVYTTTITAPPSRVLQNGGKLEKTPLYLEE
DIRPEMDIYDPTQDEKGPKEIYVWRNIIIMSLHLGLALYGITLPTCKIYIYI
WLFYLMGALGITAGARHLRSHRTYKARLPURVELIIGNTWAFQNDVFWSRDHRAH
KLFSETDADPHNSRGFFFSHVGMILLVRKHFAVKEKGSTLNLSDLRAEKLVNFORRY
KPGVLLCFILPLVPMWDETQNSLFFATLFRYALGLNVTWLVNSAAMHYGVRPY
DKTINPRNIIIVSLGAGEGPHNYHTPPDYASAEYRWHINFTTFIDCMAAIGLAY
DRKKVSKAAILARIKRGESYKSG"
6926
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="g"
8243
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="g"
8646
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="g"
10153
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="g"
10213
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="c"
10329
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="t"
13565
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="t"
13638
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="c"
13757
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="a"
14047
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="t"
14578
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="g"
14662
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="y"
15001
variation /gene="SCD"
```

```
variation /note="compared to sequence in variant B"
/replace="g"
15101
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="t"
15202
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="a"
15284
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="c"
15491
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="a"
16524
variation /gene="SCD"
/note="compared to sequence in variant B"
/replace="t"
ORIGIN
Query Match 12.8%; Score 500; DB 4; Length 17088;
Best Local Similarity 64.6%; Pred. No. 3.8e-111;
Matches 1208; Conservative 0; Mismatches 485; Indels 177; Gaps 24;
QY 2164 AAGCAAAATTAGACATTGCAGAGAGAACCTGAAGGGGGTTCAGACCAGGTACAGATTCTCT 2223
DB 1 AAGCCAAATTAGACATTGCAG--GAGAGCCTGGAGAGAGCAACACCTGCCAGACTTCTCT 59
QY 2224 GCC---ACATGCCAAGTACTTCTCAGGCATGACTGGATGAGCTGTCCACATCTGAAATC 2279
DB 60 CCCCAACATGCTGAGGGCTTCTGAGCGGTGA-TGAGACACTGTCTCTATCTAATAAAC 118
QY 2280 ATCCAGTCTTGTTCAGAACTTTTCACCCGGACAGGAGCCAGGACTGGAAATCAGTCTCC 2339
DB 119 ATCTAGTCTCATTTCAAAACTTTTTCAGTGGAGTCAAGACTGGAATGAGGTGA 178
QY 2340 TGGTCACCTGGCCAGAGAGTTGGCCCTTGACCCAGTGGCCCAACAAAGAGCTCT 2399
DB 179 TGGTGACCGGCCAGGAGGCTGGCC-AGACCCCAAGACCTGTGGCCTTCTAGGAGCTCT 237
QY 2400 TAGTCTACTCCAGGAAATCCAGGTGCTTCTCTCTGGAAGTGAATCATTTGGCGCA 2459
DB 238 TAATCATCTCCAGGCAATCCAGGTGCTCTCTCC--AGAACTAGTCTTGGCACA 295
QY 2460 GCACTCCGTATTTCTCTCTTCCAGGGGAGGATCCTAGGGCAGTATTTGGGAAAGAC 2519
DB 296 AAACCTCTCACTTCTCTCTTCCCA-----AATGATGTATTGTGGGAGAC 343
QY 2520 ATGGGCATGGAGGACACCGGTGAATCATAGCCTGCTGCTGTGAGCT---CTCAT 2575
DB 344 AGGGGTGTGTTGGAGCCCGGGAGAAATGCACGCTGCTGTGTTCCCGCTTATCCCAT 403
QY 2576 GGTAAAGGCTCTTACACACACGAAAGATGGGGGCACAGACAGTACAGTAGGCTCAGA 2635
DB 404 GGTAAAGGCTCCACACAGCCCCCGGAAAGGCAAGACTGTGAGCCACAAAGTAGGTAGA 463
QY 2636 GCATCTAGGACCCGAGGGCAATATGTCCTGAGCAGGATTAAGAGCTTGGGCTCTCAT 2695
DB 464 ACATCTAGGAGACCAGG-----GATTCAGCAGGAGTAAGA-CCTGGGCTCCATC 514
QY 2696 ATGGTGTCTTGGGCTCAACTCCAGCTCCGTCACCTACTCTGCTGCTGACCAT-GGGC 2754
DB 515 ATTACATTCTGGACTCAGTACTAGTCCCATCATTTGATTTGGCTCTGTGACCATAAAGC 574
QY 2755 AAGTTATTCCATCTCTCCATATCTCTTTCCTC---ACTTTTAAATGAATAATGGGTA 2811
DB 575 AAGTTATTATTCTCTCCATAGCTCACTTCTCTCATCTTTTAAAGGGAATAAGTACTA 634
QY 2812 CCCACCTCCAGGGTACACA--GAGAGGGCTTACAGAAACGATTTGTGATTTGGCTTGC 2869
```

Db 635 CCTACCTCACAGGGCCACAAGTTAAGATCAAAATGAATCTCTGTGGATTGCTGTG- 693
QY 2870 AGTAATAATTCAATACCTTGCAGCTATTCTTATTCACATCCAGCCCTTTCGCTGCTG 2929
Db 694 -----CAATACCTTTGGCTACTCTTCTGCTGCACTTCAAGTACTTTCTCTGCTT 743
QY 2930 CTGGGTGAAACACATGTCAGTGTTC-----TGACGGTTTCCCAAGAAAGATTC 2981
Db 744 CTGAGTGAAGACAGCTCTCTCTTCTCCAGAAATGCTTTTCCCAAGAAAGCAATTC 803
QY 2982 CAAATATACACCTCCAGCTCTGAGATCTCCAAAACATCCCGACGCAATCTCTGGAGGC 3041
Db 804 CAAACACAGCCCGAGCTTCT-----AATCTCCAAAACATCCCGACGCAATCTCTGGC 859
QY 3042 CGGGCTTTGGGATGGGACTCCCGCCCGGCTC-----TGACGGTTTCCCAAGAAAGATTC 3074
Db 860 ATGGGCACAGGACAGACAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 919
QY 3075 -----CTGACAGGATGGTGGCGGAGGACACACACAGCAGCTGTGTGTG 3126
Db 920 ACCATTGGCTGAACCTCTTACGGCGGTGCAATCTACATCAGCGCATCTCTGTGTG 979
QY 3127 CGGCGGAGTCCCG-----TGCGGTCCCG 3150
Db 980 TCGCGGCTTCCAGCTGCTGCTGCGGAGGGGAGACCCGAAACCTGTTTCCAGCAGTCCCG 1039
QY 3151 GTGAGCAGCGCTGT 3206
Db 1040 GGTTAGCGTGCCTGCTGT 1099
QY 3207 -----GCCCCCGCGCTGCTGCGGAGGGGAGCGGGGCTTCCC 3244
Db 1100 TCCTCCCGCCCTPACCCCGCATCCCGGCGCTGCGGACAGAGGAGGGGCTTGGC 1159
QY 3245 GCGTCCCTCCAGCTCCAGATCTTGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3304
Db 1160 CGGAGCGGGCCCGGGGCGGGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1219
QY 3305 GACGGGAAGACGGGAGGAGATTTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1279
Db 1220 AGCGGGAAGGCTGGGT 1279
QY 3365 CTTGAGAACTTCCCGATGCTCCAGCCCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 3424
Db 1280 CCGAGAACTTCCCGATGCTCCAGCCCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1334
QY 3425 GCTAGGC-----TCCGCGCCCGCAGCCCAACCGGCTCCCGAGCGCTTCCAGAGAAA 3480
Db 1335 GGTGTGACGATCCCGGCGCATCTCCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1394
QY 3481 GCTCCGACGGGATGCTCCGCGAGAGCCCGAGCGGGTGGAGAGAGAGAGAGAGAGAGAG 3540
Db 1395 CAGAGCG--GAGAGGTGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1449
QY 3541 AGAAACAG 3600
Db 1450 AGAAAG 1509
QY 3601 GCCAATGCAACGCGAG 3660
Db 1510 GCCAATGCAACGCGAG 1569
QY 3661 GTTTACAGAGCTTATAGCATTTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
Db 1570 GTCTACAG 1628
QY 3721 TGT 3780
Db 1629 GCCAAGCGCGGT 1686
QY 3781 TCTGTCTCTCTCCCGCTCCCGCTTACCTCCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 3840

Db 1687 TCGGGCTCTGTCT 1746
QY 3841 TCGCACCTTTGCCCTGCTTGCAGCGGATTAAGAGGGGCTGAGAAATACCGGACACGCT 3900
Db 1747 GCGCACCTTTGCCCTGCTTGTGCAACGAATAAAGAGGTCTGAGGAAATACCGGACACAGT 1806
QY 3901 CACCGGTGC 3910
Db 1807 CACCCCTGC 1816

RESULT 11

AL157935

LOCUS

DEFINITION

Human DNA sequence from clone RP11-203J24 on chromosome 9, complete

ACCESSION

AL157935

VERSION

AL157935.28

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 164201)

AUTHORS

Corby, N.

TITLE

Direct Submission

JOURNAL

Submitted (09-JAN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

COMMENT

On Nov 30, 2001 this sequence version replaced gi:16944853. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr9 RP11-203J24 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm

VRCOR: pBACe3.6

This sequence is the entire insert of clone RP11-203J24 The true left end of clone RP11-379C10 is at 163899 in this sequence. The true right end of clone RP11-228B15 is at 66809 in this sequence.

FEATURES

Location/Qualifiers

1..164201

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="9"

/clone="RP11-203J24"

/clone_lib="RP11-11.1"

55900

/note="Tandem repeat. Forced join. Gap size estimated to be approximately 280bp by restriction digest data."

62058..62326

/note="Sequence from overlapping clone RP11-228B15. Assembly confirmed by restriction digest data."

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature 153662..153692
/note="single clone region. Assembly confirmed by
restriction digest data."

ORIGIN

Query Match 11.8%; Score 460; DB 9; Length 164201;
Best Local Similarity 58.9%; Pred. No. 3.3e-101;
Matches 1107; Conservative 0; Mismatches 660; Indels 111; Gaps 14;

365 TATCATATTTATAGACAGGTTCTCACTCTGTGACCCAGGCTGGAGTGGGACAA 424
5030 TTTTCTTTCTTTGAGACAGGCTTGTCTGTGACCCAGGCTGGAAACCCAGTGGACAA 5089

425 TCATAGCTCACTGTAACTCAAACTCCCTGGGCTCAAGTGTCTCTACTACCTCAGCCCTCCA 484
5090 TCTCAGCTCACTACAACTCTGCTCTGCTGGTTCAGCAATTCCTCTGCTCAGCCCTCT 5149

485 GAGTGTAGGACTACAGGACACACAGCCATACCTGGCTAAATTTTTTTTTTAATTTTC 544
5150 GAGTGTAGGACTACAGGCTGTGACCCAGGCTGGCTTTCTTTGTTGTTTGTGAGACAGG 5269

545 ATTTTATGTATCAATTTCT-----TCTTTTTTTGTTGTTGTTTGTGATAGG 595
5210 GAGTGGGTTTCACTAATGTTGGCCAGGCTGGCTTTCTTTGTTGTTTGTGAGACAGG 5269

596 GTCTCACTTTGTACCCAGGCTGGAGGCTAGTGGCATGTGACAGCTGA--GCAGCCTTG 653
5270 GTCTTACTCTCTTTTCCAGGCTAGAGTGGAGTGGTGAATCATGGCTCACTGACAGCTCA 5329

654 ACTTCTGGGCTCAAGTGTATCTCTGCTGCTCAGGCTCCCAAGTGTGGGACTACAAACA 713
5330 ACCTCTAGGCTCAAGCAATCTCCACCTCGGCTTCCCAAGTGTGGAAATACAGACA 5389

714 CGTGTACCATGC--CTGCTCATATTTTCTTCTGAAACAGGCTATCAGTCTGTGC 770
5390 TGAGCTACCAATCTTTTTTTTTTTTTTTTTTTTTTTTGTGACACAGGCTTTGCTCTGTGC 5449

771 CCAGGCTGGAGTACAGTGGGTAATAATAGTCACTGCAGCCTCCCTCCTGGGCTCAAG 830
5450 CCAGGCTGGATTCAGTGGGACACACCTCGCTCACTGCACACCTCTGCTCTGGGCTAAG 5509

831 CAATCCGCTGGCTCAGCATCTCTGAGTGTGGGACTACAGGCTTGTGACCCAGGCTCA 890
5510 CAATCTCCTGCTCAGCATCTCTGAGTGTGGGACTACAGGCTTGTGACCCAGGCTG 5569

891 GCTAAGTTTAAATGATTTTGTGTATAGGAGGCTCTGTGTATGTTGCTCAGGCTGT 950
5570 GCTAATTTTT-----GTATTTTATAGATGGGGTTTCAGCATGTGTGGCCAGACTGG 5623

951 ATTTTATTTGTGAGCAAGGCTCACTATGTTGGCATGATCCCGCCAGCTCCACCTCCC 1010
5624 TCTCAAACTCTTGA-----CCTCAATATGATCTGCCCGCTGGCTCTCC 5667

1011 AAGTGTCTCATCTATCTCTTATAGTCACTGTGACAGCATTTAGGTTGTTTCACTTT 1070
5668 AAGTGTCTGGAATTAAGGCTATGAGCCTACGACCCGCGCCACTATTCTTTATTTTCACTTA 5727

1071 TTGACCATTTATGAATAATCTCCAGTGAATTTTATGTATATCATTTGTTGGCATATGT 1130
5728 CAGATGAGGGAACAGACCTTGAGGTTGAAGTCAATTTGCTTTGATGACACAGGAAGT 5787

1131 TTTTCAATTTCTGTGGGTTTATATCTAGGAGTGGAAATTTGCTGGATCCCGGGTAATTTTG 1190
5788 CAGCCAGGGGCTGGGATTCGTGCCAGGTAGGCTGTCTCTCACATTTTCTTTCTTT 5847

1191 ACAGGACAGTTTACGGGAGAAACCTTGGGAAATGAAGCATGTTTGAATAACAGCAA 1250
5848 TCACATTTTTTTTTTCTGTGAGATGGAGTTTCACTCTT---CCTGCCAGGCTGGAGTGC 5904

1251 GAGTGCAGGGGTTTTTGGAGTTTTTATTTTATTTCTGTGTGACAAATGTGCAATTTGATG 1310
5905 GAGACAGCCTCATCAACATGTTGAAACCCCAATGGTGGATCTTAGCTCACTGTTTCTT 5964

1311 AAGATCAAGTTTACTAAGTGAAGAGTGAAGTAAAGGCTGGAAATAGGGGCTTCAGAGT 1370
5965 TTTTGTAGGCTCGT--CAAGTGAAGCAATGAGAGTGGAGGAAGCAAAAGCAATCTGTAA 6022

1371 AAAATCATGAAGCACACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATAA 1430
6023 CTAGTTGTGATCAATTAATTTGTAACACAAATGTTTTTTTCAATTTAGTGTGCCATTTAG 6082

1431 AAAATCAAAATTTTTTTTTTTTTTTTGGAGAAAGTCTTGTCTCTTTCACTTGGCTGGAG 1490
6083 TTTTTTCTATTTATTTATTTTGTGATGGAGTCTTGTCTGTGT--TGCCAGGCTGGAA 6141

1491 GGAGTGTGTGATCTCAGCTCACTGCAACTTTTTCGCTCCCGGGTTTCAAGCAATTTCTCT 1550
6142 TGCATGGCCGATCTCGGCTCACTGCAACTCCACTCTCTGGGTTTCAAGTATTTCTCTCT 6201

1551 GCTTCAGCCTCCCAAGTAGCTGGGACTTACAGGCACCTCCACCACATGCCAGCTGATTTTT 1610
6202 GCCTTAGCCTCCCAAGTAGCTGGGACTTACAGGTGCAATGCCACCATGCCAGCTAATTTTT 6261

1611 GTATTTTGTAGTGAAGTGGGATTTCACTTTTGTGGCAAGCTGGTCTCAAACT----- 1663
6262 GTATTTTGTAGTGAAGTGGGATTTGCGCAATGTTGGCCAGGATGGGCTCAAACTCTTGACC 6321

1664 ----TTTTGCTGTCAATAATTTGTTAACTATTGTTCTTCTTGTGCTAGGTAGGGCCCCCAG 1719
6322 TCAGGTGATCCACCCGCTTGGCCTCCCAAGTGTAGGATTTATAGGCGGTGAGCCACTGC 6381

1720 ACCAAAAAATAAATCTTTAGAAATCCAAATCACTGTGTGTG-----GTTTCACCA 1768
6382 ACCCGCGCTCTCAATTTTCCACACATTTTGTGTAGCACATCTTCCCTTAGCACAT 6441

1769 CTGTCACTTGTAGAACACAGTGTGACAGGCTCTCAGGCTCAGAGTGTAGGTGTCTCTGTGCA 1828
6442 TTTTCTAAGCTTATTTCCCAAGAAAGAAACACAGCAATCTACAAGCCAGTGTCTATT 6501

1829 AAGAGAAATAGATGAATATTTCT-----GTTTGTGAGTGTGTGAGTGTGTGAGTGT 1853
6502 GAAGAAATGACCTCAAGACTTTTGAATGTGCACTGTGAACAGTGTGAGTGTGAGAC 6561

1854 -----CCGGCCAGGCTGTGGCTCATGCTGTATATCCAGCACATTTTGGAGGCTCA 1905
6562 TAAGACAGGTAGACCCAGGCTGTGGCTTACGCTGTATATCCAGCACATTTTGGAGGCTG 6621

1906 AGCATGTGGATCACTGTAGGTGAGGATTTCAAAACAGCTGGCCACACATGTTGTAACCC 1965
6622 AGGTGAGCCGATCACCTGAGGTGGGAGTTCAAGACCAACCTGACCAACGTGGAGAAACC 6681

1966 CCGTCTCTACTAAAAAATCAABAAATTAGCTAAGTGTGGTGGGATGCTGTATATCCCA 2025
6682 CCGTCTCTACTAAAAATCTA--AATTAGCAGGTGTGGTGGCATGCTGTATATCCCA 6739

2026 GCTACTTGGAGGCTGAGCAGGAGAAATTTCTTGA--CCCGGAGGCTGAGGTTGCAAGTGT 2084
6740 GCTACTTGGAGGCTGAGCAGGAGAAATGCTTTGAACCCAGGCTGGGAGGTGTGTGTTG 6799

2085 AGCCAGATCACACCACTGCACTCAGGCTGGG--GGAGAGCGGAGACTTCTCTCAAAA 2143
6800 AGCCAGATCGCGCAATTTGAGCTCTAGCTTGGGCAAGAGTGAAGTCTCCGCTCAAAA 6859

2144 AAAAAAACAAGAA 2161
6860 AAAAAAAAAAAAAA 6877

RESULT 12
AX357255
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

AX357255
Sequence 1 from Patent WO0162954.
AX357255
AX357255.1 GI:18674409
Homo sapiens (human)

617 bp
DNA
linear
PAT 13-FEB-2002

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Brownlie,A.J., Hayden,M.R., Attie,A.D., Ntambi,J.M.,
Gray-Keller,M.P. and Miyazaki,M.
TITLE Methods and compositions using stearyl-coa desaturase to identify
triglyceride reducing therapeutic agents
JOURNAL Patent: WO 0162954-A 1 30-AUG-2001;
Xenon Genetics Inc. (CA); WISCONSIN ALUMNI RESEARCH FOUNDATION
(US); University of British Columbia (CA)
FEATURES
source Location/Qualifiers
1..617
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 11.5%; Score 450.4; DB 6; Length 617;
Best Local Similarity 99.8%; Pred. No. 3e-99;
Matches 451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3459 CCCAGCCCCCTCCAGAGAGAAAGCTCCCGACGGGGATCGCGGACAGCCAGCGCGG 3518
Db 4 CCCCCCCCCCTCCAGAGAGAAAGCTCCCGACGGGGATCGCGGACAGCCAGCGCGG 63
QY 3519 GTTGGAAGAGAGCTGAGAGAGAGAAACACAGAGGGGAGCGAGGAGCTGGCGGCAG 3578
Db 64 GTTGGAAGAGAGCTGAGAGAGAGAAACACAGAGGGGAGCGAGGAGCTGGCGGCAG 123
QY 3579 AGGGAACAGCAGATTGCGCGAGCCATGCGACAGCGGAGAGCAGAGTGGACCAATTTC 3638
Db 124 AGGGAACAGCAGATTGCGCGAGCCATGCGACAGCGGAGAGTGGACCAATTTC 183
QY 3639 CTTTCGCCAATACGAGCGCGAGTTTACAGAGCCCTCATTTAGCATTTCCCGAGCGAG 3698
Db 184 CTTTCGCCAATACGAGCGCGAGTTTACAGAGCCCTCATTTAGCATTTCCCGAGCGAG 243
QY 3699 GCGAGGGGACAGCGCGGGTGGTGGTGTGCGAGCATCCCGCGGCCCTGCT 3758
Db 244 GCGAGGGGACAGCGCGGGTGGTGGTGTGCGAGCATCCCGCGGCCCTGCT 303
QY 3759 GCGGTGCGCGGAGCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3818
Db 304 GCGGTGCGCGGAGCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 363
QY 3819 ACCGCCCGCGCAGTCACTCTCGACATTGCGCCCTGCTTGGCAGCGATATAAGGGG 3878
Db 364 ACCGCCCGCGCAGTCACTCTCGACATTGCGCCCTGCTTGGCAGCGATATAAGGGG 423
QY 3879 CTGAGGAATACCGACAGCGGTTCACCGGTGC 3910
Db 424 CTGAGGAATACCGACAGCGGTTCACCGGTGC 455

```

```

RESULT 13
AC119725
LOCUS AC119725
DEFINITION Homo sapiens chromosome 3 clone RP11-759D18, WORKING DRAFT
SEQUENCE, 36 unordered pieces.
AC119725
ACCESSION AC119725
VERSION AC119725.2 GI:29293997
KEYWORDS HTG, HTGS_PHASE1, HTGS_DRAFT.
SOURCE Homo sapiens (human)

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 200288)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

```

```

Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devilla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,C., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabili,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hugues,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulsegad,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., Mcleod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okuwonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Picken,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Schreier,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodtgren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 200288)
Worley,K.C.
Direct Submission
Submitted (01-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 200288)
Worley,K.C.
Direct Submission
Submitted (27-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 27, 2003 this sequence version replaced gi:20376806.
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Drafting Center Code: BCM
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HCIR
Center clone name: RP11-759D18
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 169013 bases at least Q40
Consensus quality: 174569 bases at least Q30
Consensus quality: 178019 bases at least Q20
Estimated insert size: 177488; sum-of-ctngs estimation
Quality coverage: 3x in Q20 bases; sum-of-ctngs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces

```

* is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 1266: contig of 1266 bp in length
1267 1366: gap of unknown length
1367 1468: contig of 1082 bp in length
1468 2448: gap of unknown length
2449 2548: gap of unknown length
2549 4153: contig of 1605 bp in length
4154 4253: gap of unknown length
4254 5318: contig of 1065 bp in length
5319 5418: gap of unknown length
5419 6987: contig of 1569 bp in length
6988 7088: gap of unknown length
7089 8369: contig of 1282 bp in length
8370 8469: gap of unknown length
8470 9488: contig of 1019 bp in length
9489 9588: gap of unknown length
9589 11671: contig of 2083 bp in length
11672 11771: gap of unknown length
11772 13295: contig of 1524 bp in length
13296 13395: gap of unknown length
13396 14422: contig of 1027 bp in length
14423 14522: gap of unknown length
14523 15982: contig of 1460 bp in length
15983 16082: gap of unknown length
16083 18344: contig of 2262 bp in length
18345 18444: gap of unknown length
18445 20933: contig of 1649 bp in length
20934 20994: gap of unknown length
20995 21347: contig of 1154 bp in length
21348 21447: gap of unknown length
21448 24953: contig of 3506 bp in length
24954 25053: gap of unknown length
25054 27672: contig of 2619 bp in length
27673 27772: gap of unknown length
27773 29866: contig of 2094 bp in length
29867 29966: gap of unknown length
29967 31985: contig of 2019 bp in length
31986 32085: gap of unknown length
32086 34334: contig of 2249 bp in length
34335 34434: gap of unknown length
34435 37766: contig of 3332 bp in length
37767 37866: gap of unknown length
37867 40539: contig of 2673 bp in length
40540 40639: gap of unknown length
40640 44375: contig of 3736 bp in length
44376 44475: gap of unknown length
44476 50326: contig of 5851 bp in length
50327 50426: gap of unknown length
50427 53739: contig of 3313 bp in length
53740 53839: gap of unknown length
53840 61829: contig of 7390 bp in length
61830 61929: gap of unknown length
61930 68906: contig of 6977 bp in length
68907 78680: contig of 9674 bp in length
78681 84331: contig of 10651 bp in length
84332 89531: gap of unknown length
89532 100730: contig of 1119 bp in length
100731 100830: gap of unknown length
100831 112190: contig of 11360 bp in length
112191 112290: gap of unknown length
112291 122606: contig of 10316 bp in length
122607 135164: contig of 12458 bp in length
135165 135264: gap of unknown length
135265 147115: contig of 11851 bp in length
147116 159369: contig of 12154 bp in length
159370
  
```

```

* 159370 159469: gap of unknown length
* 159470 173326: contig of 13857 bp in length
* 173327 173426: gap of unknown length
* 173427 200288: contig of 26862 bp in length.

FEATURES
      Location/Qualifiers
          1..200288
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="3"
            /clone="RP11-759D18"

ORIGIN

Query Match      11.4%; Score 447.2; DB 2; Length 200288;
Best Local Similarity 58.5%; Pred. No. 4.6e-98;
Matches 1099; Conservative 0; Mismatches 668; Indels 111; Gaps 14;

QY 365 TATCATATTATTAGACAGAGTTCTCACTCTGTCAACCCAGGCTGGAGTGCAGTGGGACAAA 424
Db 103464 TTTCTTCTTTTGTGACAGAGGTCTTGCTCTGTCAACCCAGGCTGGAGTGCAGTGGGACAAA 103523

QY 425 TCATAGCTCACTGTAACTTCAAACTCTGGGCTCAAGTGATCCTACTACTCAGCCTCCA 484
Db 103524 TCTCAGGTCACTACAACCTCTGCTCTCTGGGTTCAGGCAATTCCTCTGCTCAGCCTCCT 103583

QY 485 GAGTAGCTAGGACTACAGGCACACAGCAGCAPACCTGGCTAAATTTTTTTTTTAAATTTTC 544
Db 103584 GAGTAGCTAGGACTACAGGCAGTGTCCCAACACAGAGATTAATTTTTTATATTTTAAAT 103643

QY 545 ATTTTATGTATTCATTTCT-----TTCTTTTGTGCTGTTGTTGTTGAGATAGG 595
Db 103644 GAGATGGGGTTTCACTATGTTGGCCAGGCTGGCTTTCTTTGTTGTTTGTGAGACAGG 103703

QY 596 GTCTCATTGTTTACCAGGCTGGAGGCGAGTGGCATGGTGACAGCTGA--GCAGCCTTG 653
Db 103704 GTCTTACTCTCTTTTCCAGGCTAGAGTGCAGTGGTGCAATCATGCTCAGTCACGCTCA 103763

QY 654 ACTTCTGGGCTCAAGTGAATCTCTGCTCAGCCTCCCAAGTACGAGTGGGACTACAAACA 713
Db 103764 ACCTCTAGGCTCAAGCAATCTCCCACTGGGCTCCCAAGTGTGGAATACAGACA 103823

QY 714 CGTGTCACCATGC---CTGGCTGATATTTTTTTTCTTCAACAGGATATCATCTGTTGC 770
Db 103824 TGAGCTACCATCTTTT-----TTTTTTTTTTTTTTTTTTTGTGAGACAGAGTCTTGCTGTGTC 103883

QY 771 CCAGCTGGAGTACAGTGGCGTAAATATAGTCACTGCAGCTCCCTCTCTGGGCTCAAG 830
Db 103884 CCAGCTGGATTGCAGTGGGCAACCTCGGCTCACTGCAACCTCTGCTCTCTGGTCAAG 103943

QY 831 CAATCCGCTGCCTCAGCATCTCAGTAGCTGGAGCTACAGGCTGTGCGACCCAGGCCCA 890
Db 103944 CAATCTCGCTCCCTCAGCATCTCAGTAGCTGGAGCTACAGGCGATGCGCCACCATGCTG 104003

QY 891 GCTAAGTTTTAAAAAATGATTTTTTGTATAGAGAGGTCTTGCTATGTTGCTCAGGCTGT 950
Db 104004 GCTAATTTTT-----GTATTTTGTAGATGGGGTTTCAGCATGTTGCCAGACTGG 104057

QY 951 ATTTTATGTTGAGACAAGTCTCAATATGTTGCCATGATCCGCCACCTCCACTTCCC 1010
Db 104058 TCTCAAACTCTGA-----CCTCAAAATGATCTGCCGCTCGGCTCCC 104101

QY 1011 AAGTGTCTCATCTTATCTGTTTCATTAGTCAAGTGCAGACATTTAGGTTGTTTCCACTTT 1070
Db 104102 AAGTGTCTGGAATTTACAGGCATGAGGCATGAGCCCTGCAACCCGCGCACTATTCTTATTTTCACTTA 104161

QY 1071 TTGACCATTTATGAATAATACTCCAGTGAATATTCAATGATACATTTGTGTTGGGCATATGT 1130
Db 104162 CAGATGAGGGAAACCAAGACCTCGAGGTGAGTCAATTCGCTTTTGTATGACACAGGAAGT 104221

QY 1131 TTTTCAATTTCTGTTGGGTTTATATCTTAGGAGTGGAAATGCTGTGGATCCCGGGTAATATTTTG 1190
Db 104222 CAGCCAGGGGCTGGGATTCGTGCCAGGTAGGCTCTGCTCTCACATTTTCTTTCTTTT 104281
  
```

RESULT 14					
AL731547					
LOCUS	AL731547	175842 bp	DNA	linear	PRI 27-SEP-2002
DEFINITION	Human DNA sequence from clone RP11-565H3 on chromosome 10,				
ACCESSION	Complete sequence.				
VERSION	AL731547	AC073586			
KEYWORDS	AL731547..9	GI:23380932			
SOURCE	HTG.				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 175842)				
AUTHORS	Wray, P.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-SEP-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:				
	humany@wanger.ac.uk Clone requests: clonerequest@sanger.ac.uk				
	On Sep 30, 2002 this sequence version replaced at: 21694732				
COMMENT					

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

These sequences were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi, EMBL; Sw, SWISSPROT; Tr, TREMBL; Mp, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/Help/Chr10>
RP11-655H3 is from the library RP11-11.2 constructed by the group of Peter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACe3.6.

OPTATA

Query Match	11.1%	Score 433	DB 9	Length 175842
Best Local Similarity	58.3%	Pred. No. 1.3e-94		
Matches 1071		Conservative	0	Mismatches 670
			Indels	96
			Gaps	14

[illegible]

422 CAATCATAGCTCACTGTAACCTCAAACCTCCTGGGCTCAAGTGATCCTACTACCTCAGCCT 481

	DB	146760	CATCATAGCTCAGTACAAACAACAACCTCCTGGGCTCAAGGATCCTCCACTCAGCCT	146819
	QY	482	CCAGAGTAGCTAGGACTACAGGCACACACAGCCATACCTGGCTAATTTTTTTTAAATT	541
	DB	146820	CCTGAGTAGCTGGGACTACAGGCACACACCTGCGCACCTGGCTAATTTTAATTT	146872
	QY	542	TTCATTTTATGTATTCATTTCTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGA	601
	DB	146873	-----TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	146910
	QY	602	CTTTTGTTACCAGGCTGGAGGCGCAGTGGCATGGTGACAGCTGA--GCAGCCCTTGACTGCC	659
	DB	146911	CTCCATCACCCAGGCTGGAGAGCAGTGGGACAAAGTTGGCTCACTGCAATCTCTGTCTCC	146970
	QY	660	TGGGCTCAAGTGATCCTCCTGGCTCAGCCTCCCAAGTAGCTGGGACTCAAAACAGTGTC	719
	DB	146971	TGGGTTCGAAGGATTTCTCCTGCTCAGCCTCCGAGTAGCTGGGACTCAGGCATGTGCC	147030
	QY	720	ACCATGCGCTGGCTGATAATT---TTTTTCTTGAAAACAGGGTATCACTCTGTGTGCCAGG	775
	DB	147031	ACCAAGCCTAGCTAATTTTGTAATTTTAGTAGATGGGGTTCTCCATGTGGCCAGG	147090
	QY	776	CTGAGTAGACGTGGGTAATTAATAGCTCACTGACGCTCCCGCTCTGGCTCAAGCAATC	835
	DB	147091	CTGGCTCGA-----ACTCTGACCTCAGGTGGTC	147120
	QY	836	CGCTGGGCTCAGCATCCTGAGTAGCTGGGACTACAGGCTTCTGGC--ACGAGGCCAGCT	893
	DB	147121	CACCCACTCGACCTCCCACAGTGGTGGGATTCAGGCTTGAGCCCGGCCAACCTGGCT	147180
	QY	894	AAGTTTTAAAAATGATTTTGGTATAGAGGAGTCTTGCTATGTGTGTCAGGCTGTATT	953
	DB	147181	AAATTTTTTAAAGAAAAATTTTGTAGAGA-CCGGCTTGTCTTCTCACCAGTGGAGT	147239
	QY	954	TTTATGTGTGAGCA-----AGGTCTCACTATGTGCAATGATCCCCCACTCCAC	1005
	DB	147240	TCAGTTATGCCATCATAGCTCACTGTAAGTCGAAATTCCTGGGCTAAAGCAATCCTCT	147299
	QY	1006	TTCCCAAAGTGCCTCATCTTATCTCTTCATTAGTCAGTAGACAGACATTTAGGTGTGTTCC	1065
	DB	147300	GCGTCACCTTCCCAAAATGCTGGGATTCAGGCATGAGCCACTGTACTGTGCCCAGTTTG	147359
	QY	1066	ACTTTTGCACCATATGCAATAATCTCCAGTGAATATCATGTATACATTTGTGTGGCCA	1125
	DB	147360	ACTTTCTAGCTCCACTTCTCGAGACTCCCGAGCTCAAGTAAAGTTCTTTCTCAACTCT	147419
	QY	1126	TATGTTTTTCATTTCTGTGTGGGTT-TATATCTAGAGTGGAAATGTGCGATCCCGGGTAAT	1184
	DB	147420	TTGGTCTTGGACAAGCAAAACTCTTCTGCA TGGGTCTTCTAGCTCTTACCTTCAT	147479
	QY	1185	ATTTTGCAGGCGAGGTTGAGGGGAAGAAAACCTGGGAAAATGAAGCATGTTTAGAAAT	1244
	DB	147480	GTTTCAGGCAACACAGATTAATAATAATTAATACTCACTCACTCATATATCCAGGATCT	147539
	QY	1245	CAGCAAGAGTGCAGGGGTTTTTCGGAGTTTTATTTTATATTTCTGTGACAAAATGTCAGT	1304
	DB	147540	CAGATCTTGCCTTAAACTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	147599
	QY	1305	TTGATGAAGATACAAGTTATATAAGTGAGAGTGAAGTAATTAAGGCTGGAAATAGGGCGTT	1364
	DB	147600	GACTCCTCCCTCTCCCTCTTTTACATTTTTTTCCTATTGATACATAATATTTTACATGTTT	147659
	QY	1365	CAGAGTAAATCATGAAGCACTTTTGAATACCAAAATTAAGGAGCT-----TGCGTGTA	1418
	DB	147660	ATGGGGTACATGATTTTGTCTACATGCATGAATATGTAGAGGTTGAACATCTCAATC	147719
	QY	1419	ACAAAATAATAAAAAATCAATTTTTTTTTTTTTTTTTTTTTTTTGTGAGAAAGAGTCTTGCTCTTCA	1478
	DB	147720	CCATTACCTTGAGTATTTATCTTTTTTTTTTTTTTTTTTTTTTTTGTGAGTGGASTCTTGCTCTGTTG	147779
	QY	1479	CCCTGGCTGGAGGCGAGTGGTGTGATCTCAGCTCACTGCAACTTTCCGCTCCCGGTTCA	1538

Db	147780	CTCAGGCTGGAGTGCACAAATGSCAAATCTCAGTTCACTGCAACTCTGCTCCACAGTTCA	147839
QY	1539	AGCAATTCCTCGCTTCAGCCCTCCAGTAGCTGGGACTACAGGCACITTCACCACTGCG	1598
Db	147840	AGTGAATTCCTCGCTTCAGCCCTCCAGTAGCTGGGACTACAGGCACITTCACCACTGCG	147899
QY	1599	CAGCTGAATTTTGTATTTTATTTAGTAGAGATGGGATTTCACTTTGTGTGGCCAAGCTGTCTC	1658
Db	147900	CGGCTAAATTTTGTATTTTATTTAGTAGAGATGGGATTTCACTTTGTGTGGCCAAGCTGTCTC	147959
QY	1659	AAACTTTTGTCTCATATTTCTTGAATACTATTTCTTTTGTCTG	1714
Db	147960	AAACTTTTGTCTCATATTTCTTGAATACTATTTCTTTTGTCTG	148019
QY	1715	CCAGACCAAAAAAATAAATCTTTAGAATCCAAATCAGTGTGTGGTTTGACCACTGTCA	1774
Db	148020	GTGAGCCACTGCACCTGGCCAAATATTTATCTTATTTCTATGTTTGGAAATATTTCAAGTC	148079
QY	1775	CTTGAGAACCAAGTGTGACAGGCGCTCAGAGGTAGAGGTGATC-----TCTGCTCGAAA	1830
Db	148080	CTCTCTTCTATCTGCTTTGAAATATATCTACACATTTCTTGCTGGCTATAGTCGGCTCCACT	148139
QY	1831	GAGAAATAGAATGAAAAATATTC-TCCGGGCCAGGCGTGGTGGCTCATCGCTGTATATCCCA	1889
Db	148140	ATCAAAACATCATATTTGTTTCTTAGCCAGGCGTGGTGGCTCATCGCTGTATATCCCA	148199
QY	1890	GCATTTGGGAGGCCAAGCCATGTGGATCACTCAGGTTCAGGAGTTCAAAACCAAGCTGG	1949
Db	148200	GCATTTGGGAGGCCAAGCCAGGAGGAGATCACTCAGGTTCGGGAGTTTCAGACCAAGCTGA	148259
QY	1950	CCAACTGGTCAAAACCCCGTCTCTACTAAATAACAAAATTAAGTGTGGTGGCG	2009
Db	148260	CCAACTGGGAGGAAACCCCGTCTCTACTAAATAACAAAATTAAGTGTGGTGGCG	148319
QY	2010	CATGCTGTATCTCCAGTACTCTGGGAGGTTGAGGAGGAGTTCCTTGAACCCGGAG	2069
Db	148320	CATGCTGTATCTCCAGTACTCTGGGAGGTTGAGGAGGAGTTCCTTGAACCCGGAG	148379
QY	2070	GCAGAGTTGCAAGTGAAGGAGATCAACCACTGCATCCAGCTGGGGGAG-GAGAGCGA	2128
Db	148380	GCAGAGTTGCAAGTGAAGGAGATCAACCACTGCATCCAGCTGGGGGAG-GAGAGCGA	148439
QY	2129	GACTTCCTCTCAAAAAACAAAAACAAAAAGAAATTA	2165
Db	148440	AATTTGATCTCAAAAAACAAAAACAAAAACGAA	148476
RESULT 15			
AC005037/c			
LOCUS	AC005037	190508 bp	DNA linear PRI 30-SEP-2000
DEFINITION	Homo sapiens BAC clone RP11-469M7 from 2, complete sequence.		
ACCESSION	AC005037		
VERSION	AC005037.2	GI:4827310	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Sulston,J.E. and Waterston,R.		
JOURNAL	Toward a complete human genome sequence		
MEDLINE	Genome Res. 8 (11), 1097-1108 (1998)		
PUBMED	99063792		
REFERENCE	2 (bases 1 to 190508)		
AUTHORS	Abbott,A. and Le,T.		
TITLE	The sequence of Homo sapiens BAC clone RP11-469M7		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 190508)		
AUTHORS	Waterston,R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-JUN-1998) Genome Sequencing Center, Washington University, School of Medicine		

RESIT.T 15

AC005037/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

—

REFERENCE

AUTHORS

TITLE	TOPIK
1. The Role of the Teacher	1. The Role of the Teacher
2. The Role of the Student	2. The Role of the Student
3. The Role of the Parent	3. The Role of the Parent
4. The Role of the Society	4. The Role of the Society
5. The Role of the Government	5. The Role of the Government
6. The Role of the Media	6. The Role of the Media
7. The Role of the Church	7. The Role of the Church
8. The Role of the Family	8. The Role of the Family
9. The Role of the Community	9. The Role of the Community
10. The Role of the Nation	10. The Role of the Nation

JOURNAL
OF
MEDICAL
EDUCATION

**MEDLINE
DISEASE**

REFERENCE

REFERENCES

TITLE

JOURNAL OF THE
 THEATRE

REFERENCES

REFERENCE AUTHORS

AUTHOR
 TITLE

FIELD JOURNAL

References

MO 63108, USA
4 (bases 1 to 190508)
Waterston, R.H.
Direct Submission
Submitted (14-MAY-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 190508)
Waterston, R.
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 14, 1999 this sequence version replaced gi:3309089.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary statistics

Center project name: H NH0469M07

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McCrerson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/qsc>

SOURCE INFORMATION:

The pC1-1 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., and de Jong, P.J. (1998). A new approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-13J8; the clone sequenced to the right is RP11-91M5. Actual start of this clone is at base position 1 of RP11-469M7; actual end is at 190508 of RP11-469M7.

FEATURES
source

```

location/qualifiers
1. .190508
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"

```

repeat_region	/rpt_family="Alu"	2392	.2562
repeat_region	/rpt_family="L1"	3284	.3315
repeat_region	/rpt_family="Alu"	3316	.3338
repeat_region	/rpt_family="(CAAA)n"	3339	.3621
repeat_region	/rpt_family="Alu"	3650	.3743
repeat_region	/rpt_family="AT_rich"	4353	.4664
repeat_region	/rpt_family="Alu"	4727	.5018
repeat_region	/rpt_family="Alu"	5545	.5585
repeat_region	/rpt_family="(CATA)n"	5586	.5873
repeat_region	/rpt_family="Alu"	5876	.6007
repeat_region	/rpt_family="Alu"	6327	.6713
repeat_region	/rpt_family="MBR21_g"	6924	.7294
repeat_region	/rpt_family="MER2_type"	8027	.8049
repeat_region	/rpt_family="AT_rich"	8050	.8354
repeat_region	/rpt_family="Alu"	8355	.8661
repeat_region	/rpt_family="Alu"	8356	.8684
repeat_region	/rpt_family="AT_rich"	9450	.9752
repeat_region	/rpt_family="Alu"	10371	.10680
repeat_region	/rpt_family="Alu"	10762	.11038
repeat_region	/rpt_family="Alu"	11039	.11109
repeat_region	/rpt_family="Alu"	11138	.11410
repeat_region	/rpt_family="Alu"	11446	.11663
repeat_region	/rpt_family="L1"	11664	.11959
repeat_region	/rpt_family="Alu"	11960	.12189
repeat_region	/rpt_family="L1"	12190	.12334
repeat_region	/rpt_family="Alu"	12335	.12360
repeat_region	/rpt_family="L1"	12464	.12514
repeat_region	/rpt_family="L1"	12515	.12822
repeat_region	/rpt_family="Alu"	12823	.12862
repeat_region	/rpt_family="L1"	12857	.13096
repeat_region	/rpt_family="L1"	13101	.13330
repeat_region	/rpt_family="MaLR"	13331	.13634
repeat_region	/rpt_family="Alu"	13635	.13692
repeat_region	/rpt_family="MaLR"	13704	.13990
repeat_region	/rpt_family="L1"	14082	.14387
repeat_region	/rpt_family="Alu"	14402	.14704
repeat_region	/rpt_family="Alu"		

[illegible]

```
FT protein_bind complement (3245..3248)
FT /*tag= bg
FT /bound moiety= "Whn"
FT protein_bind complement (3276..3279)
FT /*tag= bh
FT /bound moiety= "NF1"
FT protein_bind 3291..3294
FT /*tag= bi
FT /bound moiety= "Myc/Max"
FT protein_bind 3292..3295

Query Match 100.0%; Score 3910; DB 6; Length 4150;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAATTCATCCATTTAAATCATACAAATTTAAATGGCTTTTAGTATATTCACAGTGTGTC 60
DB 1 AGGAATTCATCCATTTAAATCATACAAATTTAAATGGCTTTTAGTATATTCACAGTGTGTC 60

QY 61 ATCCATCACAATCCATTTTAGAAGAGTTTATTACTCCAAAAATAAACCTGCAATCCCTT 120
DB 61 ATCCATCACAATCCATTTTAGAAGAGTTTATTACTCCAAAAATAAACCTGCAATCCCTT 120

QY 121 AGCCATCACCCCCCAACATCTCCATCCCTCTCCAGCCCTGGCAACCCACCAATCTAC 180
DB 121 AGCCATCACCCCCCAACATCTCCATCCCTCTCCAGCCCTGGCAACCCACCAATCTAC 180

QY 181 TTTCTGCTCTATAAATTTGCCAATTTCTGGACATTTTCAATATAATGGAACAAACAT 240
DB 181 TTTCTGCTCTATAAATTTGCCAATTTCTGGACATTTTCAATATAATGGAACAAACAT 240

QY 241 GTGAGACTTTGTGACGTGGCTTTCACATTAGCATTTCTATTTTAAAGGCTCAATGTTA 300
DB 241 GTGAGACTTTGTGACGTGGCTTTCACATTAGCATTTCTATTTTAAAGGCTCAATGTTA 300

QY 301 CAGTACTTAGCAGTACTTCATTTCTTTTATTTCTCAAAATGGTATTCACATGTGGGTAT 360
DB 301 CAGTACTTAGCAGTACTTCATTTCTTTTATTTCTCAAAATGGTATTCACATGTGGGTAT 360

QY 361 CCCATATCATATTTATAGACAGAGTTCTCACTGTCAACCAGGCTGGAGTGCAGTGGC 420
DB 361 CCCATATCATATTTATAGACAGAGTTCTCACTGTGTCAACCAGGCTGGAGTGCAGTGGC 420

QY 421 ACAATCATAGCTCACTGTAAACCTCAACCTCCGGCTCAAGTGATCCTACTACCTCAGCC 480
DB 421 ACAATCATAGCTCACTGTAAACCTCAACCTCCGGCTCAAGTGATCCTACTACCTCAGCC 480

QY 481 TCCAGAGTAGCTAGGACTACAGGCACACACAGCCATACCTGGCTAAATTTTTTTTTTAAT 540
DB 481 TCCAGAGTAGCTAGGACTACAGGCACACACAGCCATACCTGGCTAAATTTTTTTTTTAAT 540

QY 541 TTTCAATTTATGATTCATTTCTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
DB 541 TTTCAATTTATGATTCATTTCTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600

QY 601 ACTTTGTTTACCCAGCTGGAGGAGTGGCATGTGACAGCTGAGCAGCTTGACTTCCT 660
DB 601 ACTTTGTTTACCCAGCTGGAGGAGTGGCATGTGACAGCTGAGCAGCTTGACTTCCT 660

QY 661 GGGCTCAAGTGATCCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGTGTCA 720
DB 661 GGGCTCAAGTGATCCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGTGTCA 720

QY 721 CCATGCTGGCTGATATTTTTTTCTTGAAACAGGGTATCACTCTGTGTGCCAGGCTGGA 780
DB 721 CCATGCTGGCTGATATTTTTTTCTTGAAACAGGGTATCACTCTGTGTGCCAGGCTGGA 780

QY 781 GTACAGTGGGGTAAATATAGCTCACTCAGCCTCCCTCTGGCTCAAGCAATCCGCTG 840
DB 781 GTACAGTGGGGTAAATATAGCTCACTCAGCCTCCCTCTGGCTCAAGCAATCCGCTG 840

QY 841 GCCTCAGCATCCCTGAGTAGCTGGGACTACAGGCTTGTGCCACAGGCTCAGCTAAGTTT 900
DB 841 GCCTCAGCATCCCTGAGTAGCTGGGACTACAGGCTTGTGCCACAGGCTCAGCTAAGTTT 900
```

```
DB 841 GCCTCAGCATCCCTGAGTAGCTGGGACTACAGGCTTGTGCCACAGGCTCAGCTAAGTTT 900
QY 901 AAAAAATGATTTTGGTATAGAGAGGTCTGTGCTATGCTGCTCAGGCTGTATTTTATTG 960
DB 901 AAAAAATGATTTTGGTATAGAGAGGTCTGTGCTATGCTGCTCAGGCTGTATTTTATTG 960
QY 961 TTGAGACAAGTCTCACTATGTTGCCATGATCCCCACCTCCCACTTCCCAAGTGTCTCA 1020
DB 961 TTGAGACAAGTCTCACTATGTTGCCATGATCCCCACCTCCCACTTCCCAAGTGTCTCA 1020

QY 1021 TCTATCTGTTCAATAGTTCAGTTGACAGACATTTAGTGTGTTTCCACTTTTTCACCAATTA 1080
DB 1021 TCTATCTGTTCAATAGTTCAGTTGACAGACATTTAGTGTGTTTCCACTTTTTCACCAATTA 1080

QY 1081 TGAATAATACTCCAGTGAATTTATGATATACATTTATGTTGTGGGCAATGTTTTCATTTCT 1140
DB 1081 TGAATAATACTCCAGTGAATTTATGATATACATTTATGTTGTGGGCAATGTTTTCATTTCT 1140

QY 1141 GTTGGGTTTATATCTAGGAGTGAATTTGCTGGATCCGGGTAAATTTTTCACAGGACAG 1200
DB 1141 GTTGGGTTTATATCTAGGAGTGAATTTGCTGGATCCGGGTAAATTTTTCACAGGACAG 1200

QY 1201 TTCAGGGGAAGAAAACTTGGGAAAAATGAAGCATGTTTAGAAATCAGCAAGAGTGCAGGG 1260
DB 1201 TTCAGGGGAAGAAAACTTGGGAAAAATGAAGCATGTTTAGAAATCAGCAAGAGTGCAGGG 1260

QY 1261 GTTTTTTCGAGTTTATTTATTTATTTCTGTGTGACAAATGTCAGTTTGTATGAAGATACAAG 1320
DB 1261 GTTTTTTCGAGTTTATTTATTTATTTCTGTGTGACAAATGTCAGTTTGTATGAAGATACAAG 1320

QY 1321 TTATACTAAGTCAGAGTGCAGAAATTAAGGAGCTTGGCTGTAAACAAAAATAAATAAATCACA 1380
DB 1321 TTATACTAAGTCAGAGTGCAGAAATTAAGGAGCTTGGCTGTAAACAAAAATAAATAAATCACA 1380

QY 1381 AGCACTTTGAATACAAAAATTAAGGAGCTTGGCTGTAAACAAAAATAAATAAATCACA 1440
DB 1381 AGCACTTTGAATACAAAAATTAAGGAGCTTGGCTGTAAACAAAAATAAATAAATCACA 1440

QY 1441 TTTTCTGAGTCTGAGTGTGCTGTCTTTTCCCTGCTGGGAGGAGTGTG 1500
DB 1441 TTTTCTGAGTCTGAGTGTGCTGTCTTTTCCCTGCTGGGAGGAGTGTG 1500

QY 1501 TGATCTCAGCTCACTGCAACTTTGCGCTCCCGGTTCAAGCAATTTCTCTGCTCAGGCT 1560
DB 1501 TGATCTCAGCTCACTGCAACTTTGCGCTCCCGGTTCAAGCAATTTCTCTGCTCAGGCT 1560

QY 1561 CCCAAGTAGCTGGGACTACAGGCATTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
DB 1561 CCCAAGTAGCTGGGACTACAGGCATTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620

QY 1621 TAGAGATGGGATTTCACTTTGTTGCGCAAGCTGTCTCAAACTTTTGTGCTGCTGCTGCTGCTGCTG 1680
DB 1621 TAGAGATGGGATTTCACTTTGTTGCGCAAGCTGTCTCAAACTTTTGTGCTGCTGCTGCTGCTGCTG 1680

QY 1681 TTGTAACATTTGTTTCTTTTCTGCTGAGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
DB 1681 TTGTAACATTTGTTTCTTTTCTGCTGAGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740

QY 1741 AATCCAAATCAGTGTGTGTTGTTGACCACTGTCACTTTGAGAACCAAGTGTGACAGGCTG 1800
DB 1741 AATCCAAATCAGTGTGTGTTGTTGACCACTGTCACTTTGAGAACCAAGTGTGACAGGCTG 1800

QY 1801 CTGAGGATGAGGTGATCTCTGCTGAAAGAGAAATAGAAATATATCTCCGGGCT 1860
DB 1801 CTGAGGATGAGGTGATCTCTGCTGAAAGAGAAATAGAAATATATCTCCGGGCT 1860

QY 1861 AGCGTGGTGGCTCATGCTGTAATCCAGCACTTTGGGAGGCCAAGGCTGTTGATCTAC 1920
DB 1861 AGCGTGGTGGCTCATGCTGTAATCCAGCACTTTGGGAGGCCAAGGCTGTTGATCTAC 1920

QY 1921 CTGAGGTGAGGAGTTCAAAACAGGCTGCAACATGGTGAACCCCGCTCTCTACTAAAA 1980
DB 1921 CTGAGGTGAGGAGTTCAAAACAGGCTGCAACATGGTGAACCCCGCTCTCTACTAAAA 1980
```

QY	1981	ATACAAAAAATTAGCTAAAGTGTGGTCGGCGCATGCTGTATCCAGACTACTTGGGAGGGT	2040
Db	1981	ATACAAAAAATTAGCTAAAGTGTGGTCGGCATGCCGTGTAATCCAGACTACTTGGGAGGGT	2040
QY	2041	GAGGCAGAGAAATTTCTTGAAACCCGGAGGCAGAGGTTCAGTGAAGCGAGATCAACACCA	2100
Db	2041	GAGGCAGAGAAATTTCTTGAAACCCGGAGGCAGAGGTTCAGTGAAGCGAGATCAACACCA	2100
QY	2101	CTGCACTCCAGCTCGGGGAGAGAGCGAGACTTCTCTCAAAAAACAACCAAAAGA	2160
Db	2101	CTGCACTCCAGCTCGGGGAGAGAGCGAGACTTCTCTCAAAAAACAACCAAAAGA	2160
QY	2161	ATTAAAGCAAAATPAGACATTTGCAGAGAGAACCTGAAGGGGGTCAGACCGATCAGATTTC	2220
Db	2161	ATTAAAGCAAAATPAGACATTTGCAGAGAGAACCTGAAGGGGGTCAGACCGATCAGATTTC	2220
QY	2221	TGTGCCATGCCCAAGTACTTCTGAGGCATGACTGGATGAGCTGCCATCTGAAATCA	2280
Db	2221	TGTGCCATGCCCAAGTACTTCTGAGGCATGACTGGATGAGCTGCCATCTGAAATCA	2280
QY	2281	TCCAGTCTTGTTCAGAACTTTACACCCGACAGGGAGCCAGACACTGGAATGCAGTCTCT	2340
Db	2281	TCCAGTCTTGTTCAGAACTTTACACCCGACAGGGAGCCAGACACTGGAATGCAGTCTCT	2340
QY	2341	GGTCACTGGCCAGAGAGTTGSCCTTGACCTGAGACCAAGTGCACCAAAAGGAGCTGTT	2400
Db	2341	GGTCACTGGCCAGAGAGTTGSCCTTGACCTGAGACCAAGTGCACCAAAAGGAGCTGTT	2400
QY	2401	AGTCTACTCCCAAGAAATCCAGGTGCTGTCTTCTCTGGAAGTGAATCAATTCGGCGAG	2460
Db	2401	AGTCTACTCCCAAGAAATCCAGGTGCTGTCTTCTCTGGAAGTGAATCAATTCGGCGAG	2460
QY	2461	CACCTCGTATTTTCTCTCTTTCCACAGGGAAGGATCTTGGGCAAGTATTTGGGAAAGACA	2520
Db	2461	CACCTCGTATTTTCTCTCTTTCCACAGGGAAGGATCTTGGGCAAGTATTTGGGAAAGACA	2520
QY	2521	TGGGCATGAAGACACCGGGTGAATGCATAGCTGCTGTTCTGTAGACTCTCATGCTAA	2580
Db	2521	TGGGCATGAAGACACCGGGTGAATGCATAGCTGCTGTTCTGTAGACTCTCATGCTAA	2580
QY	2581	GGCTCCTACAGACCGAAAGATGGGGGCACAGGACAGATCAGTAGGGTCAGAGCATC	2640
Db	2581	GGCTCCTACAGACCGAAAGATGGGGGCACAGGACAGATCAGTAGGGTCAGAGCATC	2640
QY	2641	TCAGGACCGAGGGCAATATGTCTCTGACAGGATTAAGAGCTTGGGTCTCATATGTT	2700
Db	2641	TCAGGACCGAGGGCAATATATGTCTCTGACAGGATTAAGAGCTTGGGTCTCATATGTT	2700
QY	2701	GTTTCTGGGCTCAACTGCCAGCTCCGTCATTTACTGGTGTGACCATGGGCAAGTTA	2760
Db	2701	GTTTCTGGGCTCAACTGCCAGCTCCGTCATTTACTGGTGTGACCATGGGCAAGTTA	2760
QY	2761	TTCCATCTCTCCATATCTTTCTCTCACTTTTAAATGGAATTAAGGGTACCCACTCC	2820
Db	2761	TTCCATCTCTCCATATCTTTCTCTCACTTTTAAATGGAATTAAGGGTACCCACTCC	2820
QY	2821	CAGGTCACAGAGGGCTTACAGAAAAAGATCTTGTGAATTTGGCTTCAGTAAATATTC	2880
Db	2821	CAGGTCACAGAGGGCTTACAGAAAAAGATCTTGTGAATTTGGCTTCAGTAAATATTC	2880
QY	2881	AATACCTGCCAGCTATTCTTATTCACATCCAAAGCCCTTCGCTGCTGCTGGGTGAAA	2940
Db	2881	AATACCTGCCAGCTATTCTTATTCACATCCAAAGCCCTTCGCTGCTGCTGGGTGAAA	2940
QY	2941	CACATGTCAGTGTCTCTGACGGTTTCCAAAGAGATTCACAAATTAACACCTGCCAG	3000
Db	2941	CACATGTCAGTGTCTCTGACGGTTTCCAAAGAGATTCACAAATTAACACCTGCCAG	3000
QY	3001	TCTGAAGAAATCTCCAAAAACATCCCGCACGCACTCTGGAGGCGCGGCTTGGGATGGAC	3060
Db	3001	TCTGAAGAAATCTCCAAAAACATCCCGCACGCACTCTGGAGGCGCGGCTTGGGATGGAC	3060

QY	3061	TGCCGCGCCCGGGTCTCTGAACAGGATCGCTGCGCGCAGGCAACACACACAGCCAGCGCTG	3120
Db	3061	TGCCGCGCCCGGGTCTCTGAACAGGATCGCTGCGCGCAGGCAACACACACAGCCAGCGCTG	3120
QY	3121	TGTGTGCGGCGCGAGTCCCGTCCGGTCCGGGTGAGCAGCGCTGCTGTGTGGCGGGGC	3180
Db	3121	TGTGTGCGGCGCGAGTCCCGTCCGGTCCGGGTGAGCAGCGCTGCTGTGTGGCGGGGC	3180
QY	3181	AGAGCCATTGTTTCGACAGCGTACCCAGCCCCCGCGCTCGCCCGGAGGAGCGCGGGCT	3240
Db	3181	AGAGCCATTGTTTCGACAGCGTACCCAGCCCCCGCGCTCGCCCGGAGGAGCGCGGGCT	3240
QY	3241	TCCCGCTCCCAAGCTCCAGATCCTGGGTGGCTGCCACGTCTCCTTCCACGCGCCCTG	3300
Db	3241	TCCCGCTCCCAAGCTCCAGATCCTGGGTGGCTGCCACGTCTCCTTCCACGCGCCCTG	3300
QY	3301	GGGGAGCGGGAAGACGGACGGAGATGTTAGTGTGGCGGCCCCCGAGGGTTCCACT	3360
Db	3301	GGGGAGCGGGAAGACGGACGGAGATGTTAGTGTGGCGGCCCCCGAGGGTTCCACT	3360
QY	3361	GTTTCTCTGAGAACTTCCCGAGTCCCAACCCACCCGCTTCTCCGTGTGCCGAGGCCGGT	3420
Db	3361	GTTTCTCTGAGAACTTCCCGAGTCCCAACCCACCCGCTTCTCCGTGTGCCGAGGCCGGT	3420
QY	3421	CCTGGGTAGGCTCCGCGCCCGACCCCAACCCGGTCCCGAGCCCTTCCAGAGAGAAA	3480
Db	3421	CCTGGGTAGGCTCCGCGCCCGACCCCAACCCGGTCCCGAGCCCTTCCAGAGAGAAA	3480
QY	3481	GCTCCCGACCCGGATGCGCGGCAGAGGCCACGCGGGTGGAAAGAAAGCTGAGAAG	3540
Db	3481	GCTCCCGACCCGGATGCGCGGCAGAGGCCACGCGGGTGGAAAGAAAGCTGAGAAG	3540
QY	3541	AGAAAAGAGGGAGCGGGAGCGAGAGCTGGCGGCAGAGGGAACACAGATTCCGCCGA	3600
Db	3541	AGAAAAGAGGGAGCGGGAGCGAGAGCTGGCGGCAGAGGGAACACAGATTCCGCCGA	3600
QY	3601	GCCAAATGGCAACGGCAGGACGAGTGGCACCAATTCCTTCCGCCAATGACGAGCGGA	3660
Db	3601	GCCAAATGGCAACGGCAGGACGAGTGGCACCAATTCCTTCCGCCAATGACGAGCGGA	3660
QY	3661	GTTTACAGAAAGCTCATTAGCATTTCCCGAGGCGAGGGGACGGGACAGAGCCGGTGG	3720
Db	3661	GTTTACAGAAAGCTCATTAGCATTTCCCGAGGCGAGGGGACGGGACAGAGCCGGTGG	3720
QY	3721	TGTGTGTGCGGTGCGGACGATCCCGCGCCCTGCTGGGTTCGCCGAGCCTCGGCC	3780
Db	3721	TGTGTGTGCGGTGCGGACGATCCCGCGCCCTGCTGGGTTCGCCGAGCCTCGGCC	3780
QY	3781	TCTGTCTCTCTCCCTCCCGCCCTTACCTTCCACGCGGGACCGCCCGCGCAGTCAACTCC	3840
Db	3781	TCTGTCTCTCTCCCTCCCGCCCTTACCTTCCACGCGGGACCGCCCGCGCAGTCAACTCC	3840
QY	3841	TCGCATTGTCCTGCTTGGCAGCGGATAAAGGGGCTGAGGAAATACCGGACACGGT	3900
Db	3841	TCGCATTGTCCTGCTTGGCAGCGGATAAAGGGGCTGAGGAAATACCGGACACGGT	3900
QY	3901	CACCGGTGC	3910
Db	3901	CACCGGTGC	3910

RESULT 2
ABK94819

ABK94819
ID ABK94819 standard; DNA; 1970 BP.

AC ABK94819;

XX
DT 29-AUG-2002 (first entry)

Human stearoyl-CoA desaturase (SCD) control region.

Fatty acid regulated gene; polyunsaturated fatty acid disorder; PUFA disorder; eczema; cardiovascular disorder; hypertriglyceridaemia;

KW	dyslipidaemia; atherosclerosis; coronary artery disease; cerebrovascular disease; peripheral vascular disease; inflammation; sinusitis; asthma; pancreatitis; osteoarthritis; rheumatoid arthritis; acne; body weight disorder; obesity; cachexia; anorexia; psychiatric disorder; cancer; cystic fibrosis; pre-menstrual syndrome; diabetes; diabetic complication; genetic polymorphism; ds.
OS	Homo sapiens.
XX	WO200240666-A2.
PN	23-MAY-2002.
XX	
PD	
XX	
PF	19-NOV-2001; 2001WO-CA001632.
XX	
PR	17-NOV-2000; 2000US-0248589P.
XX	
PA	(XENO-) XENON GENETICS INC.
XX	
PI	Winther MD, Goldberg YP, Knickle LC, Haardt M, Allen SJ;
PI	Ponton A, De Antueno RJ, Jenkins DK, Nwaka SO;
XX	
DR	WPI; 2002-508327/54.
XX	
PT	Novel isolated polypeptide segment encoded by fat regulated genes, useful for diagnosing the presence of or a predisposition for a disorder involving fatty acid regulated genes in a subject.
PT	
XX	Claim 23; Fig 26; 225pp; English.
PS	
XX	
CC	The invention describes an isolated polypeptide segment (I) whose genes are fat regulated. (I) or the polynucleotide encoding it (II) are useful for diagnosing the presence of or a predisposition for a disorder involving fatty acid regulated genes in a subject. A composition containing (I) or (II) is useful for treating a disorder involving fatty acid regulated genes, where the disorder is selected from a polyunsaturated fatty acid (PUFA) disorder, eczema, cardiovascular disorders (such as hypertriglyceridaemia, dyslipidaemia, atherosclerosis, coronary artery disease, cerebrovascular disease or peripheral vascular disease), inflammation (such as sinusitis, asthma, pancreatitis, osteoarthritis, rheumatoid arthritis or acne), body weight disorders (such as obesity, cachexia or anorexia), psychiatric disorders, cancer, cystic fibrosis, pre-menstrual syndrome, diabetes, and diabetic complications. (I) or (II) is useful as research agent and materials for discovery of treatments and diagnostics for a disease, particularly human disease. (II) is useful for constructing nucleotide probes and primers, for detecting genetic polymorphism, for detecting changes in the level of expression of (II), and as a diagnostic tool. This sequence represents a control region from a gene encoding a fatty acid regulated protein
CC	
XX	
SQ	Sequence 1970 BP; 428 A; 590 C; 588 G; 364 T; 0 U; 0 Other;
Query Match	
Best Local Similarity 44.8%; Score 1751.8; DB 6; Length 1970;	
Matches 1786; Conservative 0; Mismatches 2; Indels 3; Gaps 3;	
QY	2121 GAGGCGAGACTTCCTCTCAAAAAA CAAAAA CAAAGAAATTAGCAAAATTAGACATTG 2180
Db	1 GAGGCGAGACTTCCTCTCAAAAAA CAAAAA CAAAGAAATTAGCAAAATTAGACATTG 60
QY	2181 CAGAGAACTGTAAGGGGTGCACACGTACAGATTTCTGTGCCACATGCCAAGTACT 2240
Db	61 CAGAGAACTGTAAGGGGTGCACACGTACAGATTTCTGTGCCACATGCCAAGTACT 120
QY	2241 TCTGAGGCATGACTGGATGAGCTGTCCACATCTCAAAATCATCCAGTCTTGTTCAGACTT 2300
Db	121 TCTGAGGCATGACTGGATGAGCTGTCCACATCTCAAAATCATCCAGTCTTGTTCAGACTT 180
QY	2301 TCACACGGACACGGAGCCAGGACTGGAATGCACTCTCTCGTGCACGTGGCCACAGAGATTG 2360
Db	181 TCACACGGACACGGAGCCAGGACTGGAATGCACTCTCTCGTGCACGTGGCCACAGAGATTG 240
QY	2361 GCCTTGACCTTGAGACCAAGTGGCCACAAAGGAGCTGCTTAGTCTACCTCCCAAGAAATC 2420

Db	1319	CCCAGCCCCAAACCGGGTCCCGAGCCCTTCCAGAGAGAAAGCTCCCGACGCGGGATGCC	1378
Qy	3500	GGGCAGAGGCCCCAGCGCGGGTGAAGAAGACTGAGAGAGAAACAGAGGGGAGGGGG	3559
Db	1379	GGGCAGAGGCCCCAGCGCGGGTGAAGAAGACTGAGAGAGAAACAGAGGGGAGGGGG	1438
Qy	3560	AGCGAGAGCTGGCGGCAGAGGGAAACAGCAGATTGCCCGAGCCAATGGCAAACGGCAGGA	3619
Db	1439	AGCGAGAGCTGGCGGCAGAGGGAAACAGCAGATTGCCCGAGCCAATGGCAAACGGCAGGA	1498
Qy	3620	CGAGTGGCACCAATTTCCCTTCGGCCAAATGACGAGCCGGAGTTACAGNAGCCTCATTA	3679
Db	1499	CGAGTGGCACCAAAATTTCCCTTCGGCCAAATGACGAGCCGGAGTTACAGNAGCCTCATTA	1558
Qy	3680	GCAATTTCCAGAGGCAGGGGCAGAGGCCGGGTGGTGTGTGTCGGTGTGCGCA	3739
Db	1559	GCAATTTCCAGAGGCAGGGGCAGAGGCCGGGTGGTGTGTGTCGGTGTGCGCA	1618
Qy	3740	GCATCCCGCGCCCTGCTGCGTCCGCGGAGCCTCGGCCTCTGTCTCCTCCCTCCCTCC	3799
Db	1619	GCATCCCGCGCCCTGCTGCGTCCGCGGAGCCTCGGCCTCTGTCTCCTCCCTCCCTCC	1678
Qy	3800	GCCTTTACCTCCACGCGGACCGCCCGCGGCAGTCAACTCTCGCACTTTGCCCTCGCTT	3859
Db	1679	GCCCTTACCTCCACGCGGACCGCCCGCGGCAGTCAACTCTCGCACTTTGCCCTCGCTT	1738
Qy	3860	GGCAGCGGATAAAAGGGGCTCAGGAAATACCGGCACGGTCACCCGGTGC	3910
Db	1739	GGCAGCGGATAAAAGGGGCTCAGGAAATACCGGCACGGTCACCCGGTGC	1789

RESULT 3

AAH26627

ID AAH26627 standard; DNA; 617 Bp.

XX

AC AAH26627;

DT 12-NOV-2001 (first entry)

DT 12-NOV-2001 (first entry)

Human stearoyl-CoA desaturase-1 gene promoter region.

Stearoyl-CoA desaturase-1; hSCD1; SCD1; human; promoter; screening; assay; triglyceride; very low density lipoprotein; high density lipoprotein; low density lipoprotein; cholesterol; mucous membrane; fatty acid; wax ester; anticholesterolemic; antiarteriosclerotic; therapy; ds.

—

05 Homo sapiens.

7

Key

protein_bind

5

3
4
5

misc_feature

5

protein kinase

protein_bind

— E —

protein

protein_bind

11

11

protein bind

1000

E

T protein bind

1

E 1

protein_bind

1

FF

protein_bind

FT		/*tag= h	
FT	protein_bind	/bound_moiety= "nuclear factor Y"	
FT		404..408	
FT		/*tag= i	
FT	protein_bind	/bound_moiety= "nuclear factor 1"	
FT		419..437	
FT		/*tag= j	
FT	TATA_signal	/bound_moiety= "CCAAT enhancer binding protein-alpha (C/EBP-alpha)"	
FT		467..474	
FT		/*tag= k	
XX			
PN	WO200162954-A2.		
XX			
PD	30-AUG-2001.		
XX			
PF	23-FEB-2001; 2001WO-US005855.		
XX			
PR	24-FEB-2000; 2000US-0184526P.		
PR	3-JUL-2000; 2000US-0221697P.		
PR	15-DEC-2000; 2000US-0255771P.		
XX			
PA	(XENO-) XENON GENETICS INC.		
PA	(WISC) WISCONSIN ALUMNI RES FOUND.		
PA	(UYBR-) UNIV BRITISH COLUMBIA.		
XX			
PI	Brownlie AJ, Hayden MR, Attie AD, Ntambi JM, Gray-Keller MP;		
PI	Miyazaki M;		
XX			
DR	WPI; 2001-557708/62.		
XX			
PT	Identifying an agent for treating disorders related to serum levels of		
PT	fatty acids and glycerols, by contacting screening assay with delta-9		
PT	stearoyl-CoA desaturase activity with compound and measuring the		
PT	activity.		
XX			
PS	Claim 12; Page 83; 130pp; English.		
XX			
CC	The present sequence is that of the promoter region of the human stearoyl		
CC	-CoA desaturase-1 (hSCD1) gene. The promoter structure is similar to that		
CC	of the mouse SCD1 promoter and contains conserved regulatory sequences		
CC	for the binding of several transcription factors. Like the mouse SCD1,		
CC	the human SCD1 gene is regulated by polyunsaturated fatty acids and		
CC	cholesterol at the level of gene transcription. The sterol regulatory		
CC	element binding protein (SREBP) plays a role in the transcriptional		
CC	activation of the gene. The promoter can be used to identify regulatory		
CC	elements that modulate and control SCD1 expression in humans, and to		
CC	identify regulatory proteins that are suitable targets for small molecule		
CC	intervention to modulate expression of SCD1 in humans. The invention		
CC	provides screening assays based on the role of hSCD1 in human diseases,		
CC	disorders or conditions relating to serum levels of triglyceride, very		
CC	low density lipoprotein, high density lipoprotein, low density		
CC	lipoprotein, total cholesterol, or production of secretions from mucous		
CC	membranes, monounsaturated fatty acids, wax esters, etc. Methods for		
CC	diagnosing and treating such disorders using modulators of SCD1 activity		
CC	are also provided. These disorders include various skin diseases and		
CC	diverse maladies, such as cancer and multiple sclerosis, non-insulin-		
CC	dependent diabetes mellitus, hypertension, neurological diseases, eye		
CC	diseases, and immune disorders as well as cholesterol disorders,		
CC	Lipidemia and cardiovascular disease		
XX			
SQ	Sequence 617 BP; 117 A; 211 C; 205 G; 84 T; 0 U; 0 Other;		
	Query Match	11.5%; Score 450.4; DB 5; Length 617;	
	Best Local Similarity	99.8%; Pred. No. 7.9e-83;	
	Matches 451; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	3459 CCCAGCCCTTCACAGAGAAAGCTCCGACGGGATGCCGGGCAGAGCCCAGCGGCG	3518	
Dd	4 CCCCCGCCCTTCACAGAGAAAGCTCCGACGGGATGCCGGGCAGAGCCCAGCGGCG	63	
QY	3519 GGTGAAGAAGAGCTTCAGAGAGCAACACAGAGGAGGGGACGAGGACTGGCGGCG	3578	


```
Db      64  GGTGAAGAGAGCTCAGRAGGAGAAACAGAGGGGAGGGGAGCGAGCTGGCGGCAG 123
QY      3579 AGGGAACAGAGATTGGCCGAGCCAAATGCAACGGCAGAGACGAGTGGACCAAAATTC 3638
Db      124  AGGGAACAGAGATTGGCCGAGCCAAATGCAACGGCAGAGACGAGTGGACCAAAATTC 183
QY      3639 CTTCCGGCCAATACGAGCGGAGTTTACAGAGCCCTATTAGCATTTCCCGAGGCGAGG 3698
Db      184  CTTCCGGCCAATACGAGCGGAGTTTACAGAGCCCTATTAGCATTTCCCGAGGCGAGG 243
QY      3699 GCGAGGGCAGAGCGCGGGTGTGTGGTGTGCGGTGTCGGCAGCATCCCGCGCCCTGCT 3758
Db      244  GCGAGGGCAGAGCGCGGGTGTGTGGTGTGCGGTGTCGGCAGCATCCCGCGCCCTGCT 303
QY      3759 GCGGTGCGCGAGCCTCGGCTCTGTCTCTCCCTCCGCGCTTACCTCACGCGGG 3818
Db      304  GCGGTGCGCGAGCCTCGGCTCTGTCTCTCCCTCCGCGCTTACCTCACGCGGG 363
QY      3819 ACCGCCCGCCAGTCAACTCTCGCACTTGGCCCTGCTTGGCAGCGATAAAGGGGG 3878
Db      364  ACCGCCCGCCAGTCAACTCTCGCACTTGGCCCTGCTTGGCAGCGATAAAGGGGG 423
QY      3879 CTGAGGAATACCGGACAGCGTCACCCGTTGC 3910
Db      424  CTGAGGAATACCGGACAGCGTCACCCGTTGC 455

RESULT 4
AAK83253
ID  AAK83253 standard; DNA; 8880 BP.
XX
XX  AC
XX  AAK83253;
XX
XX  DT 07-NOV-2001 (first entry)
XX
XX  DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38065.
XX
XX  KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX  KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX  OS Homo sapiens.
XX
XX  PN WC200157182-A2.
XX
XX  PD 09-AUG-2001.
XX
XX  PF 17-JAN-2001; 2001WO-US001354.
XX
XX  31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246527P.
```


DB 8193 AAAAAAAAAAAAAAGACAAA 8214
||||| ||||| ||||| ||||| |||||
RESULT 5
ABQ8096/c
ID ABQ8096 standard; cDNA; 92638 BP.
XX
AC ABQ8096;
XX
DT 18-SEP-2002 (first entry)
XX
DE Human osteoblast differentiation related cDNA SEQ ID NO 3.
XX
KW Human; osteoblast; stem cell differentiation; bone tissue deposition;
osteoporosis; osteopathic; ss.
XX
OS Homo sapiens.
XX
PN WO200250301-A2.
XX
PD 27-JUN-2002.
XX
PF 18-DEC-2001; 2001WO-US048276.
XX
PR 18-DEC-2000; 2000US-0255882P.
PR 24-APR-2001; 2001US-0285691P.
XX
PA (GENE-) GENE LOGIC INC.
PA (PROC) PROCTER & GAMBLE CO.
XX
PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI Mertz L;
XX
DR WPI; 2002-557663/59.
XX
Use of genes and their expression profiles associated with osteoblast
differentiation for screening modulators bone formation, for diagnosing
or treating e.g. osteoporosis, or as markers for the differentiation
process.
XX
Claim 1; SEQ ID NO 3; 78pp + Sequence Listing; English.
XX
The invention relates to genes and their expression profiles are used
for: (a) screening modulators of precursor stem cell differentiation into
osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
deposition of bone tissue, abnormal rate of osteoblast formation or
osteoporosis; or (c) treating or monitoring treatment of the conditions
cited in (b), or monitoring the progression of bone tissue deposition.
Specific conditions include postmenopausal osteoporosis, glucocorticoid
osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
induced abnormalities in bone formation or bone loss, conditions that
involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
or fibrous dysplasia. The present sequence is that of an osteoblast
differentiation associated cDNA marker of the invention. Note: The
sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 92638 BP; 21460 A; 24572 C; 25235 G; 21371 T; 0 U; 0 Other;
Query Match 9.6%; Score 375.4; DB 6; Length 92638;
Best Local Similarity 55.4%; Pred. No. 6.3e-67;
Matches 1046; Conservative 0; Mismatches 726; Indels 115; Gaps 12;

QY 366 ATCAATATTATTAGACAGGTTCTCAGTGTGCACCAGGTGGAGTGCGACAAT 425
57931 AGCAATTTTTTTTGAGACAGGGTCTCTCTGTCTACTCAGGTGGAGTGCGGCCAGT 57872

Db 426 CATAGTCTACTGTAACTTAACCTCAACTCCCTGGCTCAAGTCATCCTACTAGCTTCAG 485
57974 GATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

us-10-016-725-15_copy_1_3910.rng

Fri May 14 12:58:33 2004

PR	02-MAR-2000;	2000US-0186350P.	PR	02-OCT-2000;	2000US-0237040P.
PR	16-MAR-2000;	2000US-0189874P.	PR	13-OCT-2000;	2000US-0239935P.
PR	17-MAR-2000;	2000US-0190076P.	PR	13-OCT-2000;	2000US-0239937P.
PR	18-MAR-2000;	2000US-0198123P.	PR	20-OCT-2000;	2000US-0240960P.
PR	19-MAR-2000;	2000US-0205513P.	PR	20-OCT-2000;	2000US-0241221P.
PR	07-JUN-2000;	2000US-0209467P.	PR	20-OCT-2000;	2000US-0241785P.
PR	28-JUN-2000;	2000US-0214886P.	PR	20-OCT-2000;	2000US-0241786P.
PR	30-JUN-2000;	2000US-0215135P.	PR	20-OCT-2000;	2000US-0241787P.
PR	07-JUL-2000;	2000US-0216647P.	PR	20-OCT-2000;	2000US-0241808P.
PR	07-JUL-2000;	2000US-0216880P.	PR	20-OCT-2000;	2000US-0241809P.
PR	11-JUL-2000;	2000US-0217487P.	PR	20-OCT-2000;	2000US-0241826P.
PR	11-JUL-2000;	2000US-0217496P.	PR	01-NOV-2000;	2000US-0244617P.
PR	14-JUL-2000;	2000US-0218290P.	PR	08-NOV-2000;	2000US-0246474P.
PR	26-JUL-2000;	2000US-0220963P.	PR	08-NOV-2000;	2000US-0246475P.
PR	26-JUL-2000;	2000US-0220964P.	PR	08-NOV-2000;	2000US-0246476P.
PR	14-AUG-2000;	2000US-0224518P.	PR	08-NOV-2000;	2000US-0246477P.
PR	14-AUG-2000;	2000US-0224519P.	PR	08-NOV-2000;	2000US-0246478P.
PR	14-AUG-2000;	2000US-0225213P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	2000US-0225214P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246611P.
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226868P.	PR	17-NOV-2000;	2000US-0249208P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249209P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249210P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249214P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249216P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232081P.	PR	17-NOV-2000;	2000US-0249300P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250160P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251869P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0251990P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234998P.	PR	05-JAN-2001;	2000US-0259678P.
PR	26-SEP-2000;	2000US-0235484P.	XX	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000;	2000US-0235834P.	XX	Rosen CA, Barash SC, Ruben SM;	
PR	27-SEP-2000;	2000US-0235836P.	PI	WPI; 2001-483426/52.	
PR	29-SEP-2000;	2000US-0236327P.	XX	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PR	29-SEP-2000;	2000US-0236367P.	DR	useful for preventing, diagnosing and/or treating cancers and metastasis.	
PR	29-SEP-2000;	2000US-0236368P.	XX	Disclosure; SEQ ID NO 23856; 3071pp + Sequence Listing; English.	
PR	29-SEP-2000;	2000US-0236369P.	PS	AAK54951 to AAK64702 encodes the human immune/haematopoietic antigen (I)	
PR	29-SEP-2000;	2000US-0236370P.	XX		
PR	02-OCT-2000;	2000US-0236802P.	CC		
PR	02-OCT-2000;	2000US-0237037P.			
PR	02-OCT-2000;	2000US-0237038P.			
PR	02-OCT-2000;	2000US-0237039P.			

Query Match	9.4%;	Score 368.8;	DB 4;	Length 5491;
Best Local Similarity	56.8%;	Pred. No. 7.5e-66;		
Matches 1023;	Conservative	0;	Mismatches 627;	Indels 178; Gaps 10;
Qy	361	CCCATATCATATTAATTAGAGACAGGTTCTCACTCTGTCAACCAGGCTGGAGTGCAGTGGC	420	
Db	3658	CCCTAATTAATTTTTTGTGAGACGGTGTCTCACTCTGTGCGCCAAAGTTGGAGTGTAAATGAC	3599	
Qy	421	ACAATCATAGTCACTGTGTAACCTCAAACTCCTGGGCTCAAGTGATCCTACTACTACCTCAGCC	480	
Db	3598	TGAATCTTGCTCACTGCACCTTCCACCTACCGGTTCAAGTGATTCCTCCGCCCTCAGCC	3539	
Qy	481	TCCAGATGACTAGGACTACAGGCACACACAGCCATCACTGGCTCAATTTTTTTTTTTTAAAT	540	
Db	3538	TCCCAAGTACTGGACATATAGCGGCGCACCAATCAACGCTTGGCTAATTTTTTGTATT----	3483	
Qy	541	TTTTCATTTTATGATTCAATTTTTCTTCTTTTTTGTGTGTTGTTTTGTAGATAGGGTCTC	600	
Db	3482	-----TTTAGTAAAGACAAGGTTTTT	3463	
Qy	601	ACTTTGTTATCCAGGCTGGAGGCGAGTGGCATGGTGACAGCTGAGCAGCGCTTGACTTCCT	660	
Db	3462	GCCATGTTGGCCAGGCTGGT-----CTTGAACCTCAT	3432	
Qy	661	GGGCTCAAGTGTATCCTCTGTCCTCAGCGCTCCCAAGTAGCTGGGACTCAACAACGTTGCA	720	
Db	3431	GACCTCAGGTGATCGGCTCATCAGCCTCCCAAAGTGCTGGGATTACAGGCGTAGGCCA	3372	
Qy	721	CCATGCGCTGGCTGATA-----TTTTTTTTTCTGAAACAGGTTATCA	761	
Db	3371	CTGTGCTCTGCGCAATATTTATTTATTTATTACTTTTTTTTTTTTTTTAGACAGAGTCTCA	3312	
Qy	762	CTCTGTTGCCAGGCTGGAGTACAGTGGCGTAATAATAGCTCACTCAGCGCTCC-CCTCC	820	
Db	3311	CTCTGCTGCCAGGCTGGAGTGCAGTGGAGCAATCTCGGCTCACTGAAACCTCCATCTCC	3252	
Qy	821	TGGGCTCAAGCAATCCGCTGGCCTCAGCATCCTGAGTAGCTGGGACTAGAGGCTTTGGCC	880	
Db	3251	TAGGTTCAAGCTATCTCTCTGCTTCAGGCTCCCAAGTAGTTGGGATTACAGGCGGCCAAC	3192	
Qy	881	ACCAGGCCAGCTAAGTTTTTAAAAATGATTTTTTGGTATAGAGAGGCTCTGTGATGTTG	940	
Db	3191	ACCACACCCGGCTAATTTTTGTG-----ATTTTTTAGTAGAGATGAGGTTTCAACATGTTG	3138	
Qy	941	CTCAGGCTGATTTTTTTATTTGTTGAGACAAGGCTCACTATGTTGCCATGATCCCCCAAC	1000	
Db	3137	GCCAGGCGGCTCACCAATCTCTGGGCTCA-----ATGATCCACCTGCC	3094	
Qy	1001	TGCACCTTCCCAAAGTGCTCATCTTATCTGTTCATTTAGTCAGTTGACAGACATTTAGGTTG	1060	
Db	3093	TCAGCCTCCCAAGTGCTGGGATTACAGCGGTGATCCACCGCGCCCGCTGAATGTTTC	3034	
Qy	1061	TTTCCACTTTTTTGCACATATATGAATAATACTCCAGTGAATATTCATGTATACATTTGTGT	1120	

RESULT 7	
AAS42015/c	
AAS42015 standard; DNA; 9742 BP.	
IID	
XX	
XX	
AC	
XX	
DT	
17-DEC-2001 (first entry)	
Genomic sequence #331 encoding novel human enzyme polypeptide.	
XX	
DE	
XX	
KW	Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW	ligase; hyperproliferative disorder; immunodeficiency disorder;
KW	autoimmune disorder; neurological disorder; metabolic disorder;
KW	inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW	blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW	anti arthritic; nephrotropic; anticoagulant; ds.
XX	
XX	Homo sapiens.
OS	
XX	WC2000155301-A2.
PN	
XX	
PD	02-AUG-2001.
XX	
Pf	17-JAN-2001; 2001WO-US001239.
XX	
31-JAN-2000; 2000US-0179065P.	
04-FEB-2000; 2000US-0180628P.	
24-FEB-2000; 2000US-0184664P.	
02-MAR-2000; 2000US-0186350P.	
16-MAR-2000; 2000US-0189874P.	
17-MAR-2000; 2000US-0190076P.	
18-APR-2000; 2000US-0198123P.	
19-MAY-2000; 2000US-0205515P.	
07-JUN-2000; 2000US-0209467P.	
28-JUN-2000; 2000US-0214886P.	
30-JUN-2000; 2000US-0215135P.	
07-JUL-2000; 2000US-0216647P.	
07-JUL-2000; 2000US-0216880P.	
11-JUL-2000; 2000US-0217487P.	
11-JUL-2000; 2000US-0217496P.	
14-JUL-2000; 2000US-0218290P.	
26-JUL-2000; 2000US-0220963P.	
26-JUL-2000; 2000US-0220964P.	
14-AUG-2000; 2000US-0224518P.	
14-AUG-2000; 2000US-0224519P.	
14-AUG-2000; 2000US-0225213P.	
14-AUG-2000; 2000US-0225214P.	
14-AUG-2000; 2000US-0225266P.	
14-AUG-2000; 2000US-0225267P.	
14-AUG-2000; 2000US-0225268P.	
14-AUG-2000; 2000US-0225270P.	
14-AUG-2000; 2000US-0225447P.	
14-AUG-2000; 2000US-0225757P.	
14-AUG-2000; 2000US-0225758P.	
14-AUG-2000; 2000US-0225759P.	
18-AUG-2000; 2000US-0226279P.	
22-AUG-2000; 2000US-0226681P.	
22-AUG-2000; 2000US-0226868P.	
22-AUG-2000; 2000US-0227182P.	
23-AUG-2000; 2000US-0227009P.	
30-AUG-2000; 2000US-0228924P.	
01-SEP-2000; 2000US-0229287P.	
01-SEP-2000; 2000US-0229343P.	
01-SEP-2000; 2000US-0229344P.	
01-SEP-2000; 2000US-0229345P.	
05-SEP-2000; 2000US-0229509P.	
05-SEP-2000; 2000US-0229513P.	
06-SEP-2000; 2000US-0230437P.	
06-SEP-2000; 2000US-0230438P.	
08-SEP-2000; 2000US-0231242P.	
08-SEP-2000; 2000US-0231243P.	
08-SEP-2000; 2000US-0231244P.	
08-SEP-2000; 2000US-0231245P.	
08-SEP-2000; 2000US-0231246P.	
08-SEP-2000; 2000US-0231247P.	
08-SEP-2000; 2000US-0231248P.	
08-SEP-2000; 2000US-0231249P.	
08-SEP-2000; 2000US-0231250P.	
08-SEP-2000; 2000US-0231251P.	
08-SEP-2000; 2000US-0231252P.	
08-SEP-2000; 2000US-0231253P.	
08-SEP-2000; 2000US-0231254P.	
08-SEP-2000; 2000US-0231255P.	
08-SEP-2000; 2000US-0231256P.	
08-SEP-2000; 2000US-0231257P.	
08-SEP-2000; 2000US-0231258P.	
08-SEP-2000; 2000US-0231259P.	
08-SEP-2000; 2000US-0231260P.	
08-SEP-2000; 2000US-0231261P.	
08-SEP-2000; 2000US-0231262P.	
08-SEP-2000; 2000US-0231263P.	
08-SEP-2000; 2000US-0231264P.	
08-SEP-2000; 2000US-0231265P.	
08-SEP-2000; 2000US-0231266P.	
08-SEP-2000; 2000US-0231267P.	
08-SEP-2000; 2000US-0231268P.	
08-SEP-2000; 2000US-0231269P.	
08-SEP-2000; 2000US-0231270P.	
08-SEP-2000; 2000US-0231271P.	
08-SEP-2000; 2000US-0231272P.	
08-SEP-2000; 2000US-0231273P.	
08-SEP-2000; 2000US-0231274P.	
08-SEP-2000; 2000US-0231275P.	
08-SEP-2000; 2000US-0231276P.	
08-SEP-2000; 2000US-0231277P.	
08-SEP-2000; 2000US-0231278P.	
08-SEP-2000; 2000US-0231279P.	
08-SEP-2000; 2000US-0231280P.	
08-SEP-2000; 2000US-0231281P.	
08-SEP-2000; 2000US-0231282P.	
08-SEP-2000; 2000US-0231283P.	
08-SEP-2000; 2000US-0231284P.	
08-SEP-2000; 2000US-0231285P.	
08-SEP-2000; 2000US-0	

PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX
 XX WPI; 2001-465566/50.
 XX
 PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,
 PT treating neural, immune system, muscular, reproductive, pulmonary,
 PT cardiovascular, renal, proliferative disorders and cancerous diseases.
 XX
 PS Disclosure; SEQ ID NO 2141; 1180pp; English.
 XX
 CC The present invention relates to the isolation of novel human enzyme
 CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
 CC encoding them. The enzyme polypeptides of the invention may comprise the
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
 CC isomerases or ligases. The sequences of the invention are useful in the
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of
 CC disorders, including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
 CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
 CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
 CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
 CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
 CC infectious disorders (e.g. Influenza). The polynucleotides of the
 CC invention can also be used in gene therapy. AAS41685-AAS42192 represent
 CC DNA sequences encoding for the novel human enzyme polypeptides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 9742 BP; 2382 A; 2564 C; 2516 G; 2280 T; 0 U; 0 Other;
 Query Match 9.4%; Score 366.4; DB 4; Length 9742;
 Best Local Similarity 56.3%; Pred. No. 2.7e-65;
 Matches 1043; Conservative 0; Mismatches 631; Indels 178; Gaps 12;
 370 TATTATTAGACAGAGGTTCTCACTC-TGTCACCCAGGCTGAGTGCAGTGGCACAATCAT 428
 6351 TTTTCTTTGAGATGAGGTTTCACTCTTTGTTGCGGCTGAGTGCAGTGGGCGGTATCTT 6292
 429 AGCTCACTGTAACCTCAAACTCTCTGGGCTCAAGTGTATCTCACTACCTCAGCTCCAGAGT 488
 6291 GGCTCATTGCAACCTCCGGCTCCCGGTTCAAGCAATCTCTCTGCTCAGCTCTGAGT 6232
 489 AGCTAGAGCTACAGGCACACAGCCATACCTGCTAAATTTTTTTTTTAAATTTTCAATTT 548
 6231 AGCTGGGGTTACAGGCATGTGCCACACCGCCGCTAAATTTTGAATTTTAA----- 6181
 549 TATGTAATCAATTTCTTTCTTTTGTGTTGTTGTTGTTGAGATGAGGTTCTCACTTTGTT 608
 6180 -----GTAGAGGACAGGTTTCTCAATGTT 6157
 609 ACCCAGGCTGAGGAGCAGTGGCATGGTGACAGCTGACAGCGCTTGACTTCTCTGGGCTCAA 668

Db 6156 GGTCTAGGCTGGT-----CTTGAACCTCCGACCTCAG 6126
 QY 669 GTGATCCTCCTCCTCAGCCTCCCAAGTAGCTGGGACATCAAAACAGCTGTCAACCTGCT 728
 Db 6125 GTGATCCGCGCGCTCGGCTCCCAAAAGTCTGGGATTACAGCGCTGAGCCACAGCGCTC 6066
 QY 729 GGCTGATATTTTTTTTCTTTGAAACAGGTTATCACTCTGTTGCCAGGCTGGAGTACAGTG 788
 Db 6065 GGCTTTTTTTTTTTTTTTGAGATGGAGTTTGGCTCTGTCAACCAGACTGGAGTGCAGTG 6006
 QY 789 GGGTAATATAGCTCACTGACGCTC-CCCTCCTGGGCTCAAGCAATCCGCTGGCCTCAG 847
 Db 6005 GTGAGTCTCAACTTACTGCAACCTCTGGCTCTAGGTTCAAGCAATCTCATGCTCAG 5946
 QY 848 CATCTGAGTAGCTGGGACTCAGAGCTTGTGCAACAGCGCCAGCTAACTTTTAAAAAT 907
 Db 5945 CTTCCCAAGTAGATGGGACTCAGGCAATGTGCCACACCGCTGGCTAAATTTTTTATA----- 5890
 QY 908 GATTTTTTGTATAGAGGAGTCTTGTCTATGTTGCTCAGGCTGTATTTTTTATTGTTGAGAC 967
 Db 5889 TTTTTTTAGTAGAGAGCTGTTTTCGCTCTGTTGGCTAGGCTGGTCTGGAAACGCTCAGCT 5830
 QY 968 AAGGTCTCATATGTTGGCCATGATCCCCCAGCTCCACCTTCCAAAGTGTCTCATCTTATC 1027
 Db 5829 CAAG-----TGATCCGCCCACTCCGCTCCCAAGAGTGTGGGATTATA 5786
 QY 1028 TGTTCATTAGTCAGTTGACAGACATTTAGGTTGTTTCCACTTTTGGACCATATGATATA 1087
 Db 5785 GCGGTGAGCCACGGCATCTGAGTGCATGAACATTTTAAAGGGCCAGTCATCTAAAT 5726
 QY 1088 TACTCCAGTGAATATTCAATGATATACATTTGTGTGGGCATATGTTTCACTTTCTGTGGT 1147
 Db 5725 GTCTCTGGGTTAAITTCATCTGACCGCCCTTGGGTTATTCAGCAGGTTCTAGCCATAGCC 5666
 QY 1148 TTATATCTAGGAGTGGAAATGCTGGATCCCGGTATATATTTTGACAGGAGAGTTCCAGG 1207
 Db 5665 CATTTTACAGATAGGACCACTGAGTGGGATCTTAGGGAAT----- 5629
 QY 1208 GAAGAAACTTGGGAAATGAAGCATGTTAGAAATACAGCAG-AGTCAGAGGGTTTTT 1266
 Db 5628 -TAGAGAAATTAGGAACTTGTAAACGTTGACACGCCAGGAGTATAGGACACTAAGAA 5570
 QY 1267 CGAGTGTATTTTATTTATTTATTTCTGTTGAACAATGTGCAGTTTGTATGAAGATACAA 1326
 Db 5569 AGCTTTGACGCTAGGCATGGGAGCTCACACCTGTAAATCTCAGTGTCTTGGAGGCTGAG 5510
 QY 1327 TAAGTGAAGTGAAGTAATTAAGCTGGAATAGGCGGCTTCAAGTAAATCATGAGCACT 1386
 Db 5509 GCAGGAGGATCGCTTTGAGGCCAGGAGTTTGAGACCAAGCCCTGGGGAAACATAGCGAGCCCT 5450
 QY 1387 TTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAAATAAATAAATCAAAATTTTTT 1446
 Db 5449 ATCTATGAAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5390
 QY 1447 TTTTTTTTTTGAAGAGTCTTTGCTCTTTTCAACCTGGCTGGAGGGGAGTGTGATCT 1506
 Db 5389 TTTTTTTTTTGAACAGAAATCTCACTCTGTTGCCAGGCTGGAGTGCATGTTGTTATCT 5330
 QY 1507 CAGCTCAGTCAAACTTTGCGCTCCCGGTTCAAGCAATTTCTCCTGCTTCAAGCTCCCAAG 1566
 Db 5329 CAGCTCACCACAACTCCGCTCTCAGGTTCAAGTGAATCTCTCCTCAGCTCAGCTCTGAG 5270
 QY 1567 TAGCTGGGACTACAGGCACTTCCACCATGCCAGCTGATTTTTTGTATTTTTTAGTAGAGA 1626
 Db 5269 TAGTTGGGACTACAGGGGGTGCACCAATGTCCGGCTAATTTTTTGTATTTTTTAGTAGAGA 5210
 QY 1627 TGGGATTTCACTTTGTGTGGCCAAAGCTGGTCTCAAACTTTTTTGTCTGTCAATTTG----- 1680
 Db 5209 CGGGGTTTCACTATGTTGACAGGCTGGTCTCGAACTCCTGACCTTGTGATCTGCCGCC 5150
 QY 1681 -----TTGTAATAATTTGTTCC 1696
 Db 5149 TCGGCTCCCAAAAGTCTGGGATTAAGTGTGAGCCACTCGCTGGCGCTTGTGTTT 5090

QY	1697	TTTTGCTGAGTAGGCCCCCAGACCCAAATAAATCTTTAGATCCAAATCAGTGTG	1756
DB	5089	TTCTTTTACAAACACGTCGCCACCCAGATATCTTTATTTCTTTAAGAAAACCTAAGTGG	5030
QY	1757	TTGGTTTGACCACTGTCTCACTTGAGAACCAACAGTGTGACCAAGGCGCTCAGAGT---	1812
DB	5029	AGAAATGACAGACAGTTACTCAGATCATTCATATAACCGGTTAGTTGTAATTCACA	4970
QY	1813	GGTATCTCTGTCGAAAGAGAAAATAGAAATATTTCTCCGGCCAGGGCTG--GTG	1870
DB	4969	TTGGATCTGACCTCCAGGATGCACCTGTTAGAAAGAGTGGCTTGGCCCAAGGCGTGTGTG	4910
QY	1871	GCTCATGCTGTAAATCCAGCACTTTGGGAGGCCAGGCGATGTGGATCACCCTGAGTCA	1930
DB	4909	GCTCACACCTGTAATCTTAGCGCTTTGGGAGGCGAGGCGAGGCTGATTTGCTGAGTCA	4850
QY	1931	GAGTTCAAAACCAAGCTGCGCCCAACATGTGTGAACCCCGTCTCTACTTAAATAACAAA	1990
DB	4849	TAGTTAATAATAGCTGTGCACATGTTGAACCTCATCTCTACTTAAATAC-AA	4791
QY	1991	TTAGCTAAGTGTGGCGCATGCTGTAAATCCAGCTACTTGGGAGGCTGAGCGAGGAG	2050
DB	4790	TTAGCCAGCGGTGGTGCACCCACCTATATAGTCCAGCTACTTGGGGGGCTGAGCGAG	4731
QY	2051	AATTTCTTGAACCGGAGGAGGTTGCGATGAGCGAGATCACACCTGCATCTCA	2110
DB	4730	AACTGCTTGAGCGCGGAGGTGGAGGTGCAGGAGGCCAGATGTCGCCACTGCATCTA	4671
QY	2111	GCTGGGGGAGAGGAGGAGACTTCTCTCAAAAAACAAAAACAAAGAAT	2162
DB	4670	GCTTGAGGACAGAGCAAGACTCTCAAAAAAGAAAAGAAAAGT	4619
RESULT 8			
ID	AAX90847	standard; DNA; 84607 BP.	
XX	AAX90847;		
XX	13-JAN-2000	(first entry)	
DE	Human PACAP	genomic sequence.	
XX	Pituitary adenylate cyclase activating polypeptide; PACAP gene; probe;		
KW	neuropsychiatric disorder; bipolar affective disorder; BAO; BP-I; UMR;		
KW	severe bipolar affective disorder; schizoaffective disorder manic type;		
KW	SAD-M; chromosome 18p11; retinal ganglion cell; STS marker; ADCYAP1;		
KW	dinucleotide repeat polymorphism marker; W3440; detection; mutation;		
KW	clinical observation; subjective report; compound; mapping; diagnosis;		
KW	treatment; genomic sequence; ds.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	5'UTR	72668..72991	
FT		/tag= a	
FT	CDS	72993..77243	
FT		/tag= b	
FT		/product= "PACAP protein"	
FT		/note= "includes exons and introns"	
FT		77244..78651	
FT		/tag= c	
FT	3'UTR	/note= "Comprises a dinucleotide repeat polymorphism marker, W3440 and an STS marker, ADCYAP1"	
XX	WO9951762-A1.		
XX	14-OCT-1999.		
XX	02-APR-1999;	99WO-US0007401.	
XX	03-APR-1998;	98US-0080570P.	

XX	(MILL-) MILLENNIUM PHARM INC.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
----	-------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 20565 TCTCTACTAAAAATACAAAAATAGCCGGCGCTAGTGGCGGCGCTGCTGCTCCAGCT 20624
QY 2029 ACTTGGAGGCTGAGCAGAGAAATTTCTTGAACCGGAGGCGAGAGTTGCAGTGAAGC 2088
Db 20625 ACTCGGAGGCTGAGCAGAGAAATGCGTGAACCGGAGGCGAGCTTGCAGTGAAGC 20684
QY 2089 GAGATCACACCACTGACATCTCAGCCTCGGCGGAGAGAGAGAGCTTCTCTCAAAAAACA 2148
Db 20685 AAGATCGGCGCACTGCCCTCCAGCCTGGCGAGAGAGAGAGCTCGCTCTCAAAAAACA 20744
QY 2149 AAAAAAAGAAATTAAGCAAAATAGACATTGAGAGAGAA 2189
Db 20745 AAAAAAAGAAATTAAGCAAAATAGACATTGAGAGAGAA 20785

RESULT 9

AA014749
ID AA014749 standard; DNA; 12394 BP.

XX AC

XX AA014749;

XX 01-NOV-2001 (first entry)

XX Human glycogen synthase kinase 3 alpha genomic DNA.

DE Human; glycogen synthase kinase 3 alpha; antidiabetic; cytosolic;

KW antisense therapy; diabetes; hyperproliferative disorder; inflammation;

KW neurological disorder; tumour; haematopoietic disorder; infection;

XX hyperproliferative disorder; developmental disorder; ds.

OS Homo sapiens.

XX Key

FT CDS Location/Qualifiers

FT 115..11786

FT /tag= a

FT /product= "Human glycogen synthase kinase 3 alpha"

FT 115..1397

FT /tag= b

FT 398..2437

FT /tag= c

FT 2438..2625

FT /tag= d

FT 2626..5638

FT /tag= e

FT 5639..5722

FT /tag= f

FT 5723..5863

FT /tag= g

FT 5864..5974

FT /tag= h

FT 5975..7901

FT /tag= i

FT 7902..8032

FT /tag= j

FT 8033..8120

FT /tag= k

FT 8121..8227

FT /tag= l

FT 8228..9196

FT /tag= m

FT 9197..9294

FT /tag= n

FT 9295..9374

FT /tag= o

FT 9375..9470

FT /tag= p

FT 9471..9897

FT /tag= q

FT 9898..10084

FT /tag= r

FT 10085..10430

FT /tag= s

FT exon 10431..10523
FT /tag= t
FT intron 10524..11712
FT /tag= u
FT exon 11713..11786
FT /tag= v
XX WO200152865-A1.
XX 26-JUL-2001.
XX 16-JAN-2001; 2001WO-US001411.
XX 21-JAN-2000; 2000US-00488856.
XX (ISIS-) ISIS PHARM INC.
XX Monia BP, McKay R, Butler MM, Wyatt JR;
XX WPI; 2001-442247/47.
XX P-PSDB; AAE06521.
XX Antisense compound 8 to 30 nucleobases in length comprising a compound
XX that is targeted to a nucleic acid molecule encoding glycogen synthase
XX kinase 3 alpha, useful for the treatment of e.g. diabetes and
XX hyperproliferative disorders.
XX Example 15; Page 92-100; 115pp; English.
XX The invention relates to an antisense compound 8 to 30 nucleobases in
XX length targeted to a nucleic acid encoding glycogen synthase kinase 3
XX alpha. The antisense compound specifically hybridises with and inhibits
XX the expression of glycogen synthase kinase 3 alpha. The antisense
XX compound is useful for the treatment of a disease associated with
XX glycogen synthase kinase 3 alpha such as diabetes, a neurological
XX disorder, a haematopoietic disorder, a hyperproliferative disorder or a
XX developmental disorder. The antisense compounds may also be used
XX prophylactically to prevent or delay infection, inflammation or tumour
XX formation. The present genomic DNA sequence encodes human glycogen
XX synthase kinase 3 alpha protein
XX Sequence 12394 BP; 2780 A; 3043 C; 3547 G; 3024 T; 0 U; 0 Other;
XX
Query Match 8.9%; Score 349.8; DB 4; Length 12394;
Best Local Similarity 55.2%; Pred. No. 7.3e-62;
Matches 1016; Conservative 0; Mismatches 742; Indels 81; Gaps 14;
QY 370 TATTATTAGACAGGTTCTCACTCTGTCCACCCAGGCTGGAGTGCAGTGGCACAATCATA 429
Db 3604 TTTTITTTTGGATGGAGTCTTACTCTGTCCACCCAGGCTGGAGTGCAGTGGCAGATCTTG 3663
QY 430 GGTCTACTGTAACTCAAACTCCTGGGCTCAAGTGATTCCTACTACTCAGCCTCCAGAGTA 489
Db 3664 GCGCGCTACAAGCTCTGCTCCAGGTTCAAGCGATTTCTTGCCCTCAACCCACCTGAGTA 3723
QY 490 GCTAGGACTACAGGCACACACAGCCATACCTGCTAAATTTTTTTTTTTTAAATTTTCAATTT 549
Db 3724 GCTGGATTGCGAGCATGGGCCACATATGCTGGCTGAATTTTTTTTTTTTGTATT--TTTA 3781
QY 550 ATGTATTCATTTTCTTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 609
Db 3782 GTAGAGACAGGGTTTGGCCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGTAAT 3841
QY 610 CCCAGGCTGGAGGGGAGTGGGATGGTGACAGCTGAGCAGCCTTGACTTCTCTGGGCTCAAG 669
Db 3842 CCGCCCCACCTCGACCTCCCAAAGTGTGGTATTATACAGGCATGA-----GCCAGCAT 3894
QY 670 TGATCTCTCTGCTCCTCAGCCTCCCAAGTAGCTGGAGTCAAAACAGCTGTCA--CCATGCTT 728
Db 3895 ACCTAGCCCTGATTTTTCAGACAAACTGAAACTGGATTGATGTGAATCTTTTTTTT 3954
QY 729 GGCTGATATTTTCTTTTGAACAGGGTATCACTCTGTGTGCCAGGCTGGAGTACAGTG 788

PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	XX	(HUMA-) HUMAN GENOME SCI INC.	QY	1441	TTTTTTTTTTTTTTTGAGAAAGAGCTTGTCTTTCACCTTGGCTGGAGGCGAGTGGTG	1500
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB	3625	TCCTTTCTTTTTTTTTTTTCTGGAGTCTTGTCTTGTGCGCCAGGCTGGAGTGGTGGT	3684
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY	1501	TGATCTCAGCTCAGTGCACCTTTCGCTCCCGGGTTCAAGCAATTCCTCCTTCAGCCT	1560
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB	3685	CGATCTCAGCTCAGTGCACCTTTCGCTCCCGGGTTCAAGCAATTCCTCCTTCAGCCT	3744
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY	1561	CCCAAGTAGCTGGGACTACAGGCACTTCCACATGCCAGCTGATTTTTTGTATTTTAG	1620
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB	3745	CCTGAGTAGCTGGGACTACAGGCACTTCCACATGCCAGCTGATTTTTTGTATTTTAG	1620
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	QY		QY			
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	DB		DB			
PR	14-AUG-2000;	2000US-02									

PR	28-JUN-2000;	2000US-0214886P
PR	30-JUN-2000;	2000US-0215135P.
PR	07-JUL-2000;	2000US-0216647P.
PR	07-JUL-2000;	2000US-0216880P.
PR	11-JUL-2000;	2000US-0217487P.
PR	11-JUL-2000;	2000US-0217496P.
PR	11-JUL-2000;	2000US-0218290P.
PR	26-JUL-2000;	2000US-0220963P.
PR	26-JUL-2000;	2000US-0220964P.
PR	14-AUG-2000;	2000US-0224518P.
PR	14-AUG-2000;	2000US-0224519P.
PR	14-AUG-2000;	2000US-0225213P.
PR	14-AUG-2000;	2000US-0225214P.
PR	14-AUG-2000;	2000US-0225266P.
PR	14-AUG-2000;	2000US-0225267P.
PR	14-AUG-2000;	2000US-0225268P.
PR	14-AUG-2000;	2000US-0225270P.
PR	14-AUG-2000;	2000US-0225447P.
PR	14-AUG-2000;	2000US-0225757P.
PR	14-AUG-2000;	2000US-0225758P.
PR	14-AUG-2000;	2000US-0225759P.
PR	18-AUG-2000;	2000US-0226279P.
PR	22-AUG-2000;	2000US-0226681P.
PR	22-AUG-2000;	2000US-0226686P.
PR	23-AUG-2000;	2000US-0227182P.
PR	23-AUG-2000;	2000US-0227009P.
PR	30-AUG-2000;	2000US-0228924P.
PR	01-SEP-2000;	2000US-0228947P.
PR	01-SEP-2000;	2000US-0229287P.
PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	08-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	08-SEP-2000;	2000US-0232081P.
PR	12-SEP-2000;	2000US-0231968P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233365P.
PR	21-SEP-2000;	2000US-0233422P.
PR	21-SEP-2000;	2000US-0234374P.
PR	21-SEP-2000;	2000US-0234375P.
PR	25-SEP-2000;	2000US-0234597P.
PR	25-SEP-2000;	2000US-0234598P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235836P.
PR	27-SEP-2000;	2000US-0235836P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239353P.
PR	13-OCT-2000;	2000US-0239357P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-021765P.

[illegible]

Qy	2041	GAGGAGGAGAAATTTCTTGAACCCGGGAGGAGAGAGTTGCAAGTGAAGCGAGATCACACCA	2100
Db	4212	GAGGAGGAGAAATCACITTTGAACCCGGGAGGAGAGTTGCAAGTGAAGCGAGATCACACCA	4271
Qy	2101	CTGCACTCCAGCTGGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAA	2160
Db	4272	CTGCACTCCAGCTGGGGGAGAGAGCGAGAGTGGCTGTCTCAAAAAACAAAAACAAAAA	4331
Qy	2161	ATTAAGCAAAATTAGACA	2177
Db	4332	ACAAAAACAACAAAAACA	4348

RESULT 14
 ID ABL98013 standard; DNA; 29163 BP.
 XX
 AC ABL98013;
 XX
 DT 21-JUN-2002 (first entry)
 XX
 DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2665.
 XX
 KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
 KW reproductive system disorder; urinary system disorder; gene therapy;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disease; infection; cytostatic; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200155317-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001329.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 11-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.

PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	12-SEP-2000;	2000US-0232081P.
PR	12-SEP-2000;	2000US-0231968P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	25-SEP-2000;	2000US-0234998P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235836P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	29-SEP-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-0244617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.

PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227003P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229503P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 05-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 12-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235835P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 13-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239335P.
 PR 13-OCT-2000; 2000US-0239337P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249272P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251930P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCT INC.
 XX Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-465566/50.
 XX Novel polypeptides and polynucleotides useful for diagnosing, preventing,
 XX treating neural, immune system, muscular, reproductive, pulmonary,
 XX cardiovascular, renal, proliferative disorders and cancerous diseases.
 XX Disclosure; SEQ ID NO 2140; 1180pp; English.

XX The present invention relates to the isolation of novel human enzyme
 XX polypeptides (AAU2915-AAU23814), and the cDNA and genomic sequences
 XX encoding them. The enzyme polypeptides of the invention may comprise the
 XX functional classes of oxidoreductases, transferases, hydrolases, lyases,
 XX isomerases or ligases. The sequences of the invention are useful in the
 XX diagnosis, treatment, prevention and/or prognosis of a wide range of
 XX disorders including hyperproliferative disorders (e.g. cancer),
 XX immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
 XX arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
 XX disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
 XX cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
 XX (e.g. haemophilia), reproductive disorders (e.g. infertility) and
 XX infectious disorders (e.g. Influenza). The polynucleotides of the

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:19 ; Search time 281.273 Seconds
(without alignments)
7714.406 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3910

Perfect score: 3910

Sequence: 1 aggaattcattcatttaaat.....cggacacggtcacccgtgc 3910

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*

1: /cgm2_6/ptodata/2/ina/5A_COMB.seq.*

2: /cgm2_6/ptodata/2/ina/5B_COMB.seq.*

3: /cgm2_6/ptodata/2/ina/6A_COMB.seq.*

4: /cgm2_6/ptodata/2/ina/6B_COMB.seq.*

5: /cgm2_6/ptodata/2/ina/PTCUS_COMB.seq.*

6: /cgm2_6/ptodata/2/ina/backfilesl.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349.8	8.9	12394	4	US-09-488-856A-10
C 2	345	8.8	84495	4	US-09-797-906-3
C 3	322.4	8.2	3805	4	US-09-108-006C-3
C 4	321	8.2	35060	3	US-08-814-095-7
5	319.8	8.2	392000	4	US-10-027-983-11
C 6	316.6	8.1	59065	4	US-09-813-817-3
C 7	316.6	8.1	59065	4	US-09-978-197-3
8	312.4	8.0	174493	4	US-09-804-471A-3
9	312.4	8.0	174493	4	US-10-238-709-3
C 10	308.2	7.9	162450	4	US-09-345-882-1
C 11	307.6	7.9	43950	4	US-09-735-934A-3
C 12	307.6	7.9	43950	4	US-10-060-332-3
C 13	307.2	7.9	70000	4	US-09-851-896-3
C 14	304	7.8	39982	4	US-09-820-924-3
C 15	303	7.7	43950	4	US-09-735-934A-3
C 16	303	7.7	43950	4	US-10-060-332-3
C 17	301.8	7.7	246240	2	US-08-724-394A-20
C 18	301.8	7.7	246240	2	US-08-724-394A-21
C 19	301.8	7.7	246240	2	US-08-724-394A-22
C 20	299.8	7.7	39982	4	US-09-820-924-3
C 21	297.2	7.6	70000	4	US-09-851-896-3
C 22	296.6	7.6	59065	4	US-09-813-817-3
C 23	296.6	7.6	59065	4	US-09-978-197-3
C 24	291.2	7.4	83450	4	US-09-811-469-3
C 25	289.2	7.4	55298	4	US-09-491-356C-1
C 26	288	7.4	21234	4	US-09-810-671-3
C 27	288	7.4	21234	4	US-10-109-854-3

28	287.4	7.4	116592	4	US-09-818-512-3	Sequence 3, Appli
C 29	286.8	7.3	35060	3	US-08-814-095-7	Sequence 7, Appli
C 30	285.8	7.3	162450	4	US-09-345-882-1	Sequence 1, Appli
C 31	283.8	7.3	12394	4	US-09-488-856A-10	Sequence 10, Appli
C 32	282.8	7.2	3805	4	US-09-108-006C-3	Sequence 3, Appli
C 33	282.8	7.2	246240	2	US-08-724-394A-20	Sequence 20, Appli
C 34	282.8	7.2	246240	2	US-08-724-394A-21	Sequence 21, Appli
C 35	282.8	7.2	246240	2	US-08-724-394A-22	Sequence 22, Appli
C 36	280	7.2	116592	4	US-09-818-512-3	Sequence 3, Appli
C 37	275	7.0	83450	4	US-09-811-469-3	Sequence 3, Appli
C 38	275	7.0	84495	4	US-09-797-906-3	Sequence 7, Appli
C 39	270.4	6.9	72604	4	US-09-268-992-7	Sequence 7, Appli
C 40	270.4	6.9	72604	4	US-09-857-474-7	Sequence 1, Appli
C 41	269.4	6.9	23187	4	US-09-499-522-1	Sequence 6, Appli
C 42	267.2	6.8	45546	4	US-09-146-053-6	Sequence 41, Appli
C 43	265.6	6.8	21721	4	US-09-269-939A-41	Sequence 42, Appli
C 44	264	6.8	44848	4	US-09-435-739-42	Sequence 10, Appli
C 45	263	6.7	98844	4	US-09-791-211-10	

ALIGNMENTS

RESULT 1

US-09-488-856A-10
; Sequence 10, Application US/09488856A
; Patent No. 6316259
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Robert McKay
; APPLICANT: Madeline M. Butler
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF GLYCOGEN SYNTHASE KINASE 3 ALPHA EXP
; FILE REFERENCE: RTS-0115
; CURRENT APPLICATION NUMBER: US/09/488,856A
; CURRENT FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 10
; LENGTH: 12394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(397)
; NAME/KEY: CDS
; LOCATION: (2438)...(2625)
; NAME/KEY: CDS
; LOCATION: (5639)...(5722)
; NAME/KEY: CDS
; LOCATION: (5864)...(5974)
; NAME/KEY: CDS
; LOCATION: (7902)...(8032)
; NAME/KEY: CDS
; LOCATION: (8121)...(8227)
; NAME/KEY: CDS
; LOCATION: (9197)...(9294)
; NAME/KEY: CDS
; LOCATION: (9375)...(9470)
; NAME/KEY: CDS
; LOCATION: (9898)...(10084)
; NAME/KEY: CDS
; LOCATION: (10431)...(10523)
; NAME/KEY: CDS
; LOCATION: (11713)...(11786)
US-09-488-856A-10

Query Match 8.9%; Score 349.8; DB 4; Length 12394;

Best Local Similarity 55.2%; Pred. No. 5.9e-77;

Matches 1016; Conservative 0; Mismatches 742; Indels 81; Gaps 14;

QY 370 TATTATTAGACAGGTTCTCACTGTGACCCAGGCTGGAGTGCAGTGGCAATCATATA 429

DB 3604 TTTTITTTGAGATGAGTCTTACTCTGTGCCAGGCTGGAGTGCAGTGGCAATCATTTG 3663

QY 430 GCTACTGTAACTCAAACTCCTGGGCTCAAGTATCTTACTACCTCAGCCTCCAGAGTA 489
Db 3664 GCCCGCTACAACTCTGCTCCAGGTTCAAGGATCTTGTGCTCAACCACTGAGTA 3723
QY 490 GCTAGAGTACAGGCACACAGCCATACCTGGTAAATTTTTTTTTTAAATTTTCAATTT 549
Db 3724 GCTGGGATTCAGGCAATGGCCACTATGCTGGCTAAATTTTTTTTTTTTGTATT -TTTA 3781
QY 550 ATGTATTCATTTCTTCTTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 609
Db 3782 GTAGAGACAGGGTTTGTGCATGTTGGCCAGGTGTGTGTGTGTGTGTGTGTGTGTGT 3841
QY 610 CCCAGGCTGGAGGAGTGGCATGTGTGACAGCTGAGCAGCCTTGACTCTCTGGGCTCAAG 669
Db 3842 CCGCCCCACCTCGACCTCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3894
QY 670 TGATCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGTGTCA -CCATGGCT 728
Db 3895 ACCTAGCCTGATTTTCAAGACAAACTGAAACTGGATTTAGATGTGAAATCTTTTTTT 3954
QY 729 GCTGATATTTTTTCTTGAACAGGATACACTCTGTGCTGCTGCTGCTGCTGCTGCTGCT 788
Db 3955 TTTTTTTTTTTTTTTTTTGACGGAGTCTCATGCTGTCAACCAGGCTGGAGTGTGTG 4014
QY 789 GCGTAATAATAGCTCACTGCAAGCTCC -CCTCTGGGCTCAAGCAATCGCTGGGCTCAG 847
Db 4015 GGGTATCTCGGCTCACTGCAACTCCGCTCGCGGTTCAAGCGATTTCTTGTGCTCAG 4074
QY 848 CATCTGATAGTGGGACTACAGGCTGTGTGCAACAGGCCAGCTAAGTTTAAATAAT 907
Db 4075 CTTCTCTAGTAGTGGGACTACAGGCTGTGTGCAACCTCTCGGCTAATTTTTTTGTATT 4134
QY 908 GATTTTGTGTATAGGAGGTCTTCTGTATGTGTCTCAGCTGTATTTTATTTGTG - 963
Db 4135 CTAGTAGGACGGGTTTCAACGATGTAGCCAGATGTTCTCTATTTTTTTTTTTTTTT 4194
QY 964 ----AGACAAGTCTCACTATGTTGCCATGATCCGCCACCTCCACTTCCCAAGTGCTC 1019
Db 4195 TTTAAGACAGAATCTCGTTCTCTCACTAAGCTGTGGTGTGATGTGCTGCTCA 4254
QY 1020 ATCTTATCTGTTCATAGTCACTGTGACAGACATTTAGTTGTGTTCCACTTTTGACCAAT 1079
Db 4255 CTGCAACCTCTGCTCCTGGGTTCAAGGCTGCAACCTCTGCTCTGGTTCAAGCAAT 4314
QY 1080 ATGAATAATACCTCAGTGAATTCATGTATATTTGTGTGGGCATATGTTTCAATTT 1139
Db 4315 TCTTGTACTCATCCACCTGATGTGTGATTCAGGCGTGGCCACCATGCCAGCTA 4374
QY 1140 TGTGGGTTTATCTAGGAGTGAATTTGCTGGATCCCGGGTAATTTTGTACAGGCAGA 1199
Db 4375 ATTTTTTTGTATTTTAGTAGATGGGTTTGTGCACTGGTGGCCAGCTGTCTCGAAC 4434
QY 1200 GTTCAGGGAAGAAA-----ACTTGGGAATGAAGCATGTTTGAATCA 1246
Db 4435 TCTGGCTCAAGTGAITTCCTGCTTGGCTCCCAAGTGTGGGATTACAGGGGTGA 4494
QY 1247 GCAAGAGTGCAGGGTTTTTGGAGTTTTATTTATTTCTGTTCACAAATGTGCAGTTT 1306
Db 4495 GCCACATGCCAGCTGTTTTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4552
QY 1307 GATGAAGATACAAGTTATCTAAGTGAAGTGAAGTAAAGGCTGGAATAGGGGGTTC 1366
Db 4553 CATGCTGTAACTTGAATTTGAGAGCC -CGAGGCAGGAGGATTTGCTGAGACTAGA 4611
QY 1367 GAGTAAATCATGAGCACTTTGAATACAAAATTAAGAGCTTGGCTGTGTAAACAAATA 1426
Db 4612 GTTCAAAACCACTGGCCCAACATAGCCAGGTTCTTTTAAATAATAATAATAAT 4671
QY 1427 ATAAAAATCAAAATTTTTTTTTTTTTTTTTTTTGTGAGAAAGTCTTGTCTTTTCCCTGGCT 1486
Db 4672 TTTATCTATTTATTTATTTATTTATTTATTTATTTTGTGACAGAGCTGTGCGCCAGGCT 4731

QY 1487 GGAGGCAGTGTGTGATCTCAGTCTCACTGCAAACTTTTCGCTCCCGGTTTCAAGCAATTC 1546
Db 4732 GGAGTGCAGTGGCGGATCTCAGTCTCACTGCAAGCTCGCCTCTCGGGTTACGGCAATTC 4791
QY 1547 TCTGCTTCAAGCTCCCAAGTAGTGGGACTACAGGCACTTCCCAACCATGCCAGCTCA- 1605
Db 4792 TCTGCTTCAAGCTCCCGAGTAGTGGGACTACAGGCTCGCCACCATGCTTGGCTAAT 4851
QY 1606 --TTTTGTATTTTAGTAGAGTGGGATTTCACTTTTGTGGCCAGCTGCTCAAACT 1663
Db 4852 TTTTGTATTTTAGTAGAGAGAGGGTTTCAACGTTTAAACAGGATGGTCTCAATCT 4911
QY 1664 TTTTGTCTCATATAATTTGTAACTATTATTCTTTTGTCTGAGGTAGGCGCCCCAGAGCA 1723
Db 4912 CTTGAC-----TTCTGTGATCCACCACTCAGCTCCCAAGTGTGGATTAACAGGGGT 4966
QY 1724 AAAAAATAAATCTTAGAATCCAAATCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1783
Db 4967 GAGCCACCAAGCTGGGCT-----GGCCTATCTTTTAAACTTTTATTTTGGAGAA 5019
QY 1784 CACAGTGTGACAGGCTCAGGAGTAGAGGTGATCTCTGCTCGAAGAGAAATAGATG 1843
Db 5020 AAAAAATCAGAGGTGCCATTTGGCTTTTACATGTCAGCAATAAGTTGAAAAAATTTT 5079
QY 1844 AAAAAATTTCTCCGGCCCAAGGCTGTGTGCTCATGCTGTAAATCCAGCACTTTTGGAGGC 1903
Db 5080 TTTTAAGTGGGTTGGCTGGGCGGCTGCTCAGCGCTGTATCCAGCACTTTTGGAGGC 5139
QY 1904 CAAGCATGTGGATCA-----CTTGAAGTCAAGAGT 1934
Db 5140 TGAGGCTGTGGATCATGAGTCAAGGAGGCTGAGGAGTGGATCAAGGTCAGGAGA 5199
QY 1935 TCAAAACCAAGCTCGCCCAACATGTGTGAAACCCCTCTCTACTTAAATAATACAAAAATTTAG 1994
Db 5200 TCGAGACCATCTGGCTAAAGTGTGAAACCCCATCTCTACTTAAATATAC -AAAAATTTAG 5258
QY 1995 CTAAGTGTGTGGCGCATGCTGTAAATCCAGCTACTTTGGAGGTTGAGGAGGAGATTT 2054
Db 5259 CTGGGCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5318
QY 2055 TCTTGAACCCGGAGGAGGAGGTTGAGTGTGAGGAGATCAACCACTGCTGCTGCTGCTGCTGCTGCTGCT 2114
Db 5319 GCTTGAACCCGGAGGAGGAGGTTGAGTGTGAGGAGATTTGAGCCACTGCTGCTGCTGCTGCTGCTGCTGCT 5378
QY 2115 GGGGAGAGAGGAGAGCTTCTCTCAAAAAACAAAAA 2153
Db 5379 -GGCAACAGACAGAGACTCTCTCAAAAAACAAAAA 5416

RESULT 2

US-09-797-906-3/c
; Sequence 3, Application US/09797906
; Patent No. 6329188

GENERAL INFORMATION:

; APPLICANT: Zianghe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001151CIP
; CURRENT APPLICATION NUMBER: US/09/797,906
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 84495

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)...(84495)

; OTHER INFORMATION: n = A,T,C or G

US-09-797-906-3

Query Match 8.8%; Score 345; DB 4; Length 84495;
Best Local Similarity 55.0%; Pred. No. 2.6e-75;
Matches 979; Conservative 0; Mismatches 715; Indels 85; Gaps 12;

445 AAATCCCTGGGCTCAAGTGAATCTACTACCTCAGCC-TCCAGAGTAGTAGGACTACAGG 503
18461 AACTCTGGGCTGAAGCATCCACTGCTTGGGCTTCCAAAGTGGTGGGATTTACAG 18402
504 CACACAGCCATACCTGCTCAATTTTTTTTAAATTTTCAATTTATGTAATTTTC 563
18401 CATTAGTCAACCGGCTGAGCAATTTATATATATATATATATATATATATTTT 18342
564 TTTCTTTTTTG-TTGTGTGTTTTGAGATAGGGTCTCACTTTGTTACCGAGCTGGAGG 622
18341 TT 18282
623 GCAGTGGCATGGTGAAGCTGA--GCAGGCTTGAATTCCTGGGCTCAAGTGATCTCCCTG 680
18281 GCAGTGGCATGATCTCAGCTCACTGCAACTTGGACTCCCTGGGTTCCAGCAATTTCTCCTG 18222
681 CCTCAGCTCCCAAGTAGCTGGGACTACAAACACAGTGTCAACATGCTGCTGGCTGATATTTT 740
18221 CCTCAGCTCTTAAGTAGCTGGGATACAGGCACATGTGACCGGCGCCAGCTAAATTTT 18162
741 TTTTCTTGAACAGGGTATCACTCTGTGTCGCCAGGCTGGAGTACAGTGGCGTAAATATAG 800
18161 TTTTCTT-----TTTTTGTTTTTTTTTTTTGTATAGACAGGATTTTCACTATATATAG 18109
801 CTCACCTGCAGCTCCCTCCTGGGCTCAAGCAATCCGCTGCTCAGCATCTCAGTAGC 860
18108 CCAGGCGATCTCGAACTCTGAACTCAAGTATCTGCGGCTCAGCTCCCAAGTGT 18049
861 TGGGACTACAGGCTGTGGCACCAGGCGCCAGCTAAAGTTTTTAAATAATGATTTTGGTATA 920
18048 TAAGATTACAGGCATAGACCAACCATGCGCGGCTCTCTATTTTATAGTTTAAAGACACAA 17989
921 GAGGAGTCTTGTATGTTGCTCAGGCTGATTTTTTATTTTATGTTGAGACAGGCTCACTAT 980
17988 GA----GTCCCACTCTGTGTCAGGCTGGAGTGCAGTGGGCACCAATCATAGCTCACTGCA 17933
981 GTTGCCATGATCCCCCACTCCACTCCCAAGTGTCTCATTTATCTG--TTTCAATAGT 1038
17932 GCCTTAAACTCTCGGCTTAAAGCAATCTCTACTCAGCTCCGAGTAGCTACATGCAC 17873
1039 CAGTTGACAGCAATTTAGTGTGTTTCCACTTTTGGACCAATTTATGAATTAATTAATCTCAGTGA 1098
17872 ATGCCACCATCCAGCTAAATTTTATATGTTTATTAAGATGGGGTCTCACTATGTTG 17813
1099 ATATTCATGTATACATTTGTGTGGGCATATGTTTCTATTTCTGTTGGTTTATATCTAGG 1158
17812 CCCAGGCTGGTCTGGAACCTCCTGGCTTAAAGCATCTCCCATCTTGGCTTTCCAAAGTG 17753
1159 AGTGAATTTGCTGATCCCGGTATATTTTGACAGGCGAGTTCAGGGGAAGAAAAT 1218
17752 CTGGGATTACAGTGTGAGCCACCATGCTGGCCACATGATCTTTTCAAAGGTCACTT 17693
1219 TGGGAAATGAAGCATGTTTGAATAATCAGACAGATGCGAGGTTTTTGGGAGTTTTTAT 1278
17692 GACTCAGTGTG--AGATAACTGGAGAGTATATATAGTTTAAATAGGCAATTTTAA 17635
1279 TTATATCTGTTGAACAATGTCCAGTTTGTATGAAGTACAGTTTATCTAAGTGAGAGT 1338
17634 ATGACCATACACAGCCCTCACTTATGATTAATTAATCATTAATTTACGATCATTTCT 17575
1339 GAGATTAAGGCTGGAATAGGGCTTCCAGATGAAATCATGAGCACTTTGA---ATAC 1394
17574 CTATTTTATGGACATTTAGACTATTTCCAGTTCTTTGCTAATATATATATTAAGGCAC 17515
1395 CAAAATTAAGGAGCTTGGCTGTAAACAAATAATAAATAATCAAAATTTTTTTTTTTTTT 1454
17514 TGTATTAACATCTTGAACCTAAATCTTTGGCAGCATCTCTAAATTTATTTTCTTTT 17455
1455 TTGAGAAAGAGTCTTCTCTTTTCCCTGCTGGAGGCGAGTGGTGTGATCTCAGCTCAC 1514

17454 TTTTGACAGAGTCTCACTGTTGCCAGGCTCAGTGTGAGTGTGCAATCCAGCTCAC 17395
1515 TGCACATTTCCCTCCCGGTTTCAAGCAATTTCTCTGCTTCCAGCTCCCAAGTAGCTGGG 1574
17394 TGCACATTTCCACCTCCCTGGGTTCAAGTGAATTTCTCTGCTTCCAGCTCCCAAGTAGCTGGG 17335
1575 ACTACAGCACTTCCACCATGCTCCAGCTGATTTTTTGTATTTTATTTAGTAGAGATGGATTT 1634
17334 ATTACAGTGGGCGCACCATGCTGGCTAAATTTTTTGTACTTTTACAGCAGACGAGGTTT 17275
1635 C-AC-TTGTGGCCAAAGCTGGTCTCAAACTTTT----- 1666
17274 GCCCATTTTGGCCAGGCTGGTCTCAAAATTCCTGACCTCTGGGATTAACAGGCGTAGGCA 17215
1667 -----TGCTGTCAATTAATTTTGTGTAATTTTGTCTTGTCTGAGTAGGCGCCCA 1718
17214 CCACGCCGACTTAATTTATTTTCTAAATATAGATTTCTAGAAGTGAATTAAGTAAATA 17155
1719 GACCAAAAAATAAATCTTAGAATCCAAATCAGTGTGTGTTGTTGACCACTGTCACTTG 1778
17154 GTTAAAGGTATGAATTTTGTGGCTATTTTATAAAATTTGTTTATAAAGATAAAGATT 17095
1779 AGAACACACAGTGTGACCGGCTCAGAGGTAGAGTGTCTCTGCTCGAAAGAGAAATA 1838
17094 TTGATCAATTAAATATAAAAAATCTACTTTGCAAAATGGATATCTTTTCACTATCCAGTC 17035
1839 GAAT-----GAAATAATCTCCGGGCCAGGCGTGGTGGCTC 1874
17034 AAATTATATACGCGGAATGTGCTTTTAAAGATGAATAACTAGGCGCAGCGAGTGGCTG 16975
1875 ATCCCTGTAAATCCAGCACTTTTGGAGGCGCAAGCATGTGGATCACTGAGGTCAAGGT 1934
16974 ACCTCTGTAAATCCAGCACTTTAGGAGGCTGAGTGGGCGGATTAAC--AAGGTCAAGGT 16917
1935 TCAAAACAGCTGGCGCAATGCTGAAACCCGCTCTCTACTAAATAAATAAATAAATAAATAA 1994
16916 TCGAGACCATCTGGCTTAAACATGCTGAAACCCCATCTCTACTAAATAAATAAATAAATAA 16957
1995 CTAAAGTGTGTGGCGCATGCTGTAATCCAGCTTACTTGGAGGCTGAGCGAGAGAAATT 2054
16856 CCGGCTGTGTGGCGGCGCTGTAGTCCAGCTACTCTGGAGGCTGAGCGAGGAGAAATT 16797
2055 TCTTGAACCGGAGGAGAGGTTGCAAGTGAAGCGAGATCACACCTGCACTCAGCTCAGCT 2114
16796 TCTTGAACCGGAGATGAGGTTGCAAGTGAAGCGAGATTCACCACTGCACTCAGCTCAGCT 16737
2115 GGGGAGAGAGGAGAGCTTCTCTCAAAAAAACAABAAA 2153
16736 GGGCAACAGGCGCAGACTCTGATTTCAAAAAAAGAAAAA 16698

RESULT 3
US-09-108-006C-3/c
; Sequence 3, Application US/09108006C
; Patent No. 6524613
; GENERAL INFORMATION:
; APPLICANT: Steer, Clifford J.
; Kren, Betsy T.
; Bandyopadhyay, Paramita
; Roy-Chowdhury, Jayanta
; TITLE OF INVENTION: Hepatocellular Chimeraplasty
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kimeragen, Inc.
; STREET: 300 Pleasant Run
; CITY: Newtown
; STATE: PA
; COUNTRY: USA
; ZIP: 18940
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

Query Match.	8.2%;	Score 322.4;	DB 4;	Length 3805;
Best Local Similarity	68.3%;	Pred. No. 2.2e-70;		
Matches 49%;	Conservative	0;	Mismatches 216;	Indels 14; Gaps 3;
QY	1446	TTTATTTTATTTTGAAGAGTCTTGCTCTTCCCTTCCACCTGCGTGCAGGCGCAGTGTGTGATC	1505	
Db	1876	TTATGTATTATTAACAGGGTCTGTCTCTGTCAACCCAGGCTGGAGTGCAGCGCCCTGAAC	1817	
QY	1506	TCAGCTCACTGCACATTTTCGCTCCCGGGTTCAGCAAAATCTCGTCTTCAGCGCTCCCAA	1565	
Db	1816	AUTGTTCACTGCAACCTTGACCTTCTGGGCTCAAGCGATCCTCTCGCTCAGCGCTCTGTGA	1757	
QY	1566	GTAGCTGGGACTACAGGCACTTCCCAACATGCCCAGCTGATTTTTGTATTTTAGTAGAG	1625	
Db	1756	GTAGCTGAGAGCACAGGTGTGTGGCCATATGCTGGCTAATTTTGTATTTTAGTAGAG	1697	
QY	1626	ATCGGATTCACATTTGTGSCCAAGCTGGTCTCAAAC	1673	TTTTGCTGTC
Db	1696	ACAGGTCTCACATGTTGCCAGGCTGGTGTGAATCTCTGGGCTCAAGTGAATCTCTCC	1637	
QY	1674	ATAATTGTGTAACTATGTTTCTTTTGTGTAGGTAGGGCCCCCAGACCAAAAAAATAA	1733	
Db	1636	ACCTTGSCCTCCAAAGTGTGGGATTCAGGGGTGAGCCACCGGCCCGGAAAGCA	1577	
QY	1734	ATCTTAGAATCCAAATCAGTGTGTGTGTGTGACCATGTCATCTTGAACACCAAGTGTGA	1793	
Db	1576	TGTATTGAGTGTCTTGTGTGCCCAAGCAGGCGCTGTGTGGGGTGTATGAGAAATAAGA	1517	
QY	1794	CCAGGGCCTCAGGATGAGGTGATCTCTGCTCGAAGAGAAATAGAAATCAAAATAT-TC	1852	
Db	1516	TCAGAGTGCCTGCTGTGTATTCATATCTGCTGTGATGCAATGCAATTTAGAAACCTCTAAC	1457	
QY	1853	TCCGGGCCAGGCGTGGTGTCAATGCTGTAAATCCAGCACTTTGGGAGGCCAAGGCATG	1912	


```

NAME/KEY: promoter
LOCATION: 4089..22464
OTHER INFORMATION: /function= "ACHE Promotor"
OTHER INFORMATION: /standard_name= "ACHE Promotor"
FEATURE:
NAME/KEY: exon
LOCATION: 22465..22537
OTHER INFORMATION: /function= "non-translated"
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: 24090..25177
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "(translation start:
OTHER INFORMATION: 24110)"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: 25524..26009
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: 27005..27274
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: 27255..28007
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: terminator
LOCATION: 27385..27387
FEATURE:
NAME/KEY: exon
LOCATION: 28008..28129
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: terminator
LOCATION: 28129..28131
FEATURE:
NAME/KEY: exon
LOCATION: 34528..34895
OTHER INFORMATION: /function= "arsenite resistance
OTHER INFORMATION: gene"
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: 34092..34358
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: 33779..33963
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: 33493..33591

```

```

OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: 33297..33408
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: exon
LOCATION: 32959..33094
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: exon
LOCATION: 32569..32628
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 7
FEATURE:
NAME/KEY: exon
LOCATION: 32386..32468
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 8
FEATURE:
NAME/KEY: exon
LOCATION: 31894..32080
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 9
FEATURE:
NAME/KEY: exon
LOCATION: 31363..31534
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 10
FEATURE:
NAME/KEY: exon
LOCATION: 30816..31011
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 12
FEATURE:
NAME/KEY: exon
LOCATION: 30470..30626
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 13
FEATURE:
NAME/KEY: exon
LOCATION: 30187..30274
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 14
FEATURE:
NAME/KEY: exon
LOCATION: 29945..30073
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 15
FEATURE:
NAME/KEY: exon
LOCATION: 29664..29856
OTHER INFORMATION: /gene= "ARS"
OTHER INFORMATION: /number= 16
US-08-814-095-7

```

```

Query Match      8.2%; Score 321; DB 3; Length 35060;
Best Local Similarity 68.8%; Pred. No. 1.6e-69;
Matches 511; Conservative 0; Mismatches 205; Indels 27; Gaps 4;

Qy 1445 TTTTTCCTTCTTGAAGAGTCTTCTTCCCTGGCTGGAGGCACTGGTGAT 1504
      |||||
Db 9802 TCTTTTCTTTTGAGACGGAGTCTCGCTCTATCGCCAGGCTGGAGTGCAGTGGCAAT 9861
      |||||

Qy 1505 CTCAGCTCACTGCAACTTTCGCCCTCCCGGGTTCAAGCAATTCCTCTGCTTCAGCCTCCA 1564

```

```

Db      9862 CTCGGCTCACTGCAACCTCTGCCACCCAGGTTCAAGCAATTTCTCTGTGTCAGCCCTCCCA 9921
QY      1565 AGTAGCTGGGACTACAGGCACTTCCACCATGCCAGCTGATTTTGTATTTTATAGTAGA 1624
Db      9922 AATAGCTGGGATTACCAAGTCCCAACCAATATGCCAGCTAATTTTATATTTTACAGTAGA 9981
QY      1625 GATGGGATTTCACTTTGCTGGCAAGTGTCTCAAACTTTTGTCTGTCTATATTTGTGT 1684
Db      9982 GACAGGTTTCAACATGTTGGCCAGGCTGTCTCAAAATCTCTGCTCAAGTGTCTGCC 10041
QY      1685 AAC-----TATTTCTCTTTTGTCTAGGTTAGGCCCCCAGACCAAAAAATA 1732
Db      10042 CACTCAGCCTCCCAAGTGTGGGATTAAGATGTGAGCAGCTCACTCAGCAGATC 10101
QY      1733 AATCTTAGAATCCAAATCACTGTGTTGGTTTGACCACTGTCACTTGAGAACCA----- 1785
Db      10102 CCTTCCAAATTAATAATCAACAGAAACCCCAAGAAATCAATTTGTTGGCAATGCTCT 10161
QY      1786 -----CAGTGTGACCAAGGCTCAGGAGTAGAGTGATCTCTCTCGAAAGAGAAATAG 1839
Db      10162 AGTAATATGTGCATCTTGGGTGCAAGCGTCTAAATTCCTAACTGTCTGAGTTTA 10221
QY      1840 AATGAATATTTCCGGGCGAGCGGTGGTGTCTATGCTGTATCCAGCACTTTGGG 1899
Db      10222 CATTAAAGATGCTTCCGGCGAGCGGTGGTGTCTACGCTGTAAATCCAGCACTTTGGG 10281
QY      1900 AGCCAGGAGTGTGATCACTCAGTGTGAGGTTCAAAACCAAGCTGCCCAACATGGT 1959
Db      10282 AAGCCGAGGAGGTGGATCACTTCCAGTGTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTG 10341
QY      1960 GAAACCCCGTCTCTACTAAAAATACAAAAATTTAGCTAAGTGTGGTGGCGCATGCTGTGA 2019
Db      10342 GAAACCCCACTCTACTAAAAA-ATACAAAAATTAGCCAGGTGTGGTGGCGCATGCTGTGA 10400
QY      2020 ATCCAGCTACTTGGGAGGTGAGGAGGAGAGATTTCTTAACCCGGAGGAGGAGTTG 2079
Db      10401 ATCCAGCTACTTGGGAGGTGAGGAGGAGAGATTTCTTAACCCGGAGGAGGAGATTG 10460
QY      2080 CAGTAGGAGGAGATCACACCACTCACTCCAGCCTGGGGGAGA-GAGCGAGACTTCTCT 2138
Db      10461 TAGTGAGTGTGATGTCACCATGTCATCTCCATCTCTGGCAACAGGAGCAAACTCCGCT 10520
QY      2139 CAAAAAACAACAAAAACAAAAAGAA 2161
Db      10521 CAAACAAACAAACAAACAGAAA 10543

RESULT 5
US-10-027-983-11
; Sequence 11, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
; LENGTH: 392000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 137740
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 137742
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (138122)...(138221)
; OTHER INFORMATION: exon 10:intron 10

;
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 145507
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 151967
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (151967)...(1542066)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 154217
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (164037)...(164136)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (174657)...(174756)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (186224)...(186323)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (195242)...(195341)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 202703
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (202771)...(202870)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (206246)...(215602)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (218126)...(218225)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (220360)...(220459)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (222717)...(222816)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (223981)...(224080)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (227487)...(227586)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (230157)...(230256)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (232299)...(232398)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (236552)...(2366651)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (238789)...(248788)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: exon
; LOCATION: (118288)...(119101)
; OTHER INFORMATION: exon 1C
; NAME/KEY: exon:intron junction
; LOCATION: (151129)...(151130)
; OTHER INFORMATION: exon 5:intron 5
; NAME/KEY: exon:intron junction
; LOCATION: (299248)...(299249)
; OTHER INFORMATION: exon 9:intron 9
; NAME/KEY: exon:intron junction
; LOCATION: (348578)...(348579)
; OTHER INFORMATION: exon 10:intron 10

```

```

; NAME/KEY: intron
; LOCATION: (348579)...(381838)
; OTHER INFORMATION: intron 10
; NAME/KEY: intron:exon junction
; LOCATION: (386185)...(386186)
; OTHER INFORMATION: intron 11:exon 12
US-10-027-983-11

Query Match      8.2%; Score 319.8; DB 4; Length 392000;
Best Local Similarity 67.2%; Pred. No. 1.le-68;
Matches 533; Conservative 0; Mismatches 217; Indels 43; Gaps 4;

QY 1419 ACAGAAATATAAAAAATCAAAATTTTTTTTTTTTTTTTGGAGAAAGAGCTTCTCTTTTCA 1478
Db 66965 ACTAAATTTTTTTCCTACTCTACATTTCTTTTCTTTTGGACACAGATTACTCTGTCA 67024

QY 1479 CCTGTGCTGGAGGCGAGTGGTGATCTCAGTCACTGCAACTTTCGCTCCGGGTTC 1538
Db 67025 CTTAGGCTGGAGCGCAGTGGGCAATCTTGGTCACTGCAACTCCACTCTCTGGGTTC 67084

QY 1539 AGCAATCTCTGCTTTCAGCTCCCAAGTAGCTGGGACTACAGGCACTTCCCACCATGCC 1598
Db 67085 AGCAATCTCTGCTTTCAGCTCCCAAGTAGCTGGGACTACAGGCGGTAGAGCGGTGCCACCATGCC 67144

QY 1599 CAGCTGAATTTTGTATTTTATTTAGTAGAGTGGGATTTCACTTTGTTGGCCCAAGCTGCTC 1658
Db 67145 TGCGTAATTTTGTATTTTATTTAGTAGAGCAGGCTTTCAATCAAGTTGGTGGGCTGCTC 67204

QY 1659 AACT-----TTTCTGCTCATAAATTTGTAACTATTGTTCTTTTCTCTAG----- 1706
Db 67205 CAACTCCTGACCTTGGCTTCCCAAGTGTGTTATTTACTTTTCAACCTCTGTTATGCT 67264

QY 1707 -----GTAGGGCCCCCAGACCAAAAAATAAATCTTAGAATCCAAATCAAGTG 1754
Db 67265 ACCTAATTTATTTGCTGCTCTCTACACTAGATAAAGTCTATGAGGGTAGACTTCTT 67324

QY 1755 TGTGTTTGACCACTGCTCACTTGGAGNACCAGTGTGACCGGGCCTCAGGAGTAGAGG 1814
Db 67325 TGTTTTGTCCAATGCTCTATTTTCCAAATATTATTAACCGTACTGGCGGTAGTACTGCA 67384

QY 1815 TGATCTCTGCTCGAAGAG-----AAATAGAAATGAAAAATTTCTCG 1856
Db 67385 ATTATAGTTTTTGAATAAATTTAGAGTAAACAATTTCTACTGGTGAATAAATTAATG 67444

QY 1857 GGCCAGCGGTGGTGTCTATGCTCTGTAATCCAGACACTTTGGAGGCCAAGGATGTGA 1916
Db 67445 GGCGGGCGGTGATGGGTCAAGCCTGTAATCCAGACACTTTGGAGGCCGAGGCGGCAGA 67504

QY 1917 TCACCTGAGTCAAGAGTTCAAAACAGCCTGCCCAACATGGTGAACCCCGTCTACT 1976
Db 67505 TCACCTGAGTCAAGAGTTCAAGACAGCCTGGGCAACATGGTGAACCCCGTCTACT 67564

QY 1977 AAAAAATACAAAAATTAGCTTAAGTGGTGGCGCATGCTCTTAATCCAGCTACTTTGGGA 2036
Db 67565 AAAAAATAC-AAAAATTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 67623

QY 2037 GGGTGAGGAGAGAAATTTCTTGAACCCGGAGGAGAGGTTTCACTGAGGAGATCAC 2096
Db 67624 GGGTGAGGAGAGAAATTTCTTGAACCCGGAGGAGAGGTTTCACTGAGGAGGAGGAGG 67683

QY 2097 ACCACTGCACCTCAGGCTGGGGAGAGAGGAGACTTCTCTCTCAAAAAACAAAAACAA 2156
Db 67684 GCCACTGCATCCAGCTTGGCGNACAGAGCGAGACTCTGTCTCAAAATATAATAAATAA 67743

QY 2157 AAGAAATTAAGCAA 2169
Db 67744 TAAAAATTAATAA 67756

RESULT 6
US-09-813-817-3/c
; Sequence 3, Application US/09813817
; Patent No. 6340583

```

RESULT 6
US-09-813-817-3/c
; Sequence 3, Application US/09813817
; Patent No. 6340583

QY 1801 CTCAGGAGTAGAGTGATCTCTGCTCGAAGAGAAATAGAAATATTTCTCGGGCC 1860
DB 126735 GTCCTTCGAACTTGTTATTTCTCCCTGTTAATACAGTAAAAATTAACCTAAGTTGGCC 126794
QY 1861 AGGCGTGCTGCTCATGCTGTAATCCAGACACTTTGGGAGGCCAAGCATGTGATCAC 1920
DB 126795 AGGAGTGTGGCAGATGCTGTAATCCAGACACTTTGGGAGGCCAAGCATGTGATCAC 126854
QY 1921 CTGAGGTACAGAGTTCAAAAACAGGCTGGCCAAACATGCTGAAACCCCTCTCTACTAAAA 1980
DB 126855 CTGAGGTACAGAGTTGAGACAGGCTGGCCAAACATGCTGAAACCCCTCTCTACTAAAA 126914
QY 1981 ATACA----AAAAATAGCTAAGTGTGTGGCGCATGCTGTAATCCAGACTTTGGGA 2036
DB 126915 AAAAATTTGAAAAGTTAGCTGGGTGTGTGGTGTGCGCACTGTGTAGTCCAGCTATTTCGGGA 126974
QY 2037 GGGTACGAGGAGAAATTTCTTGAACCCGGAGGAGAGAGTTGCACTGAGCGAGATCAC 2096
DB 126975 GGGTACGAGGAGAAATTTCTTGAACCCGGAGGAGAGAGTTGCACTGAGCGAGATCAC 127034
QY 2097 ACCACTGCACCTCAGCTGGGGGAGAGAGAGAGTCTCTCTCAAAAAACAAAAACAA 2156
DB 127035 GCCACTGCACCTCAGCTGGGGGAGAGAGAGTCTCTCTCAAAAAACAAAAACAA 127093
QY 2157 AAGAATTAAGCAAAATAGACATTTGAGAGAGAACTTGAAGGGGTGAGACACGATACAG 2216
DB 127094 AGTTAAGTAGTACTTTGGGGCCCTATCAGATAGTGTCTCTCGAGCGAGGCTGGGCCCTTA 127153
QY 2217 TTTCTGTGCCACAT 2230
DB 127154 CTACAGTTTCACTT 127167

RESULT 9

US-10-238-709-3
; Sequence 3, Application US/10238709
; Patent No. 6680188
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164DIV
; CURRENT APPLICATION NUMBER: US/10/238,709
; CURRENT FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(174493)
; OTHER INFORMATION: n = A,T,C or G
US-10-238-709-3

Query Match 8.0%; Score 312.4; DB 4; Length 174493;
Best Local Similarity 61.7%; Pred. No. 5.2e-67;
Matches 490; Conservative 0; Mismatches 298; Indels 6; Gaps 3;

QY 1441 TTTTCTTTTCTTTTGTGAGAAAGAGCTTGTGCTTTTCACTGGCTGGTGGAGGGGAGTGGTG 1500
DB 126376 TCTCTCTTTTCTTTTCAAGACAGAGTCTTGTCTGTGACCCAGGCTGGAGTGGTGG-G 126434
QY 1501 TGAATCTCAGTCACTCAACTTTGCGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAAGCT 1560
DB 126435 CGACCTTGCTCTACTGCAACCTCCGCTCTCTGGTTCAAGAGATTTCTCTGCTCAGCTC 126494
QY 1561 CCCAAGTAGCTGGAGTACAGGCACTTCCCAACCATGCCAGGCTGATTTTGTATTTTAG 1620
DB 126495 CCCGAGTAGCTGGATACAGGTTGCCCGCCACCATGCCTGGCTAAATTTTGTATTTTAG 126554

QY 1621 TAGAGATGGGATTTTCACTTTGTTGTTGGCCAGAGCTGGTCTCAAACTTTTTCGTGTCATAATTG 1680
DB 126555 TAGAGATGGGATTTTCAACCATGTTGAAGCATTTTATAGAAAAGGGTTTAGGAGAGGAGACA 126614
QY 1681 TTGTAACCTATTGTTCTTCTTCTGAGGTAGGGCCCCCAGACCAAAAAATAAATCTTAG 1740
DB 126615 NNN 126674
QY 1741 AATCCAAATAGTGTGTTGTTGACCACTGTCTCACTTGAGAACCAACAGTGTGACAGGGC 1800
DB 126675 NNTCTCCCTCTTCTCCAAAATATGCGAGACTCTTCTGTCCCTTAGTCACTTCTCATATA 126734
QY 1801 CTCAGGAGTAGAGTGATCTCTGCTCGAAGAGAAATAGAAATATTTCTCCGGGCC 1860
DB 126735 GTCCTTCTGAACTTGTGTTTATTTCTTCTGCTGTAATACAGTAAAAATTAACCTAAGTTGGCC 126794
QY 1861 AGGCGTGTGCTCATGCTGCTGTAATCCAGACACTTTTGGGAGGCCAAGCATGTGATCAC 1920
DB 126795 AGGAGTGTGGCAGCATGCTGTAATCCAGACACTTTTGGGAGGCCGAGCGAGCATCAC 126854
QY 1921 CTGAGGTACAGAGTTCAAAAACAGGCTGGCCAAACATGTGAAACCCCTCTCTACTAAAA 1980
DB 126855 CTGAGGTACAGAGTTTGAAGACAGGCTGGCCAAACATGTGAAACCCCTCTCTACTAAAA 126914
QY 1981 ATACA----AAAAATAGCTAAGTGTGTGGCGCATGCTGTAATCCAGACTTCTGGGA 2036
DB 126915 AAAAATTTGAAAAGTTAGCTGGGTGTGTGGTGTGCGCACTGTGTAGTCCAGACTATTTCGGGA 126974
QY 2037 GGGTACGAGGAGAAATTTCTTGAACCCGGAGGAGAGAGTTGCACTGAGCGAGATCAC 2096
DB 126975 GGGTACGAGGAGAAATTTCTTGAACCCAGGAGGAGAGTTGCACTGAGCGAGATCAC 127034
QY 2097 ACCACTGCACCTCAGCTGGGGAGAGAGAGAGTCTCTCTCAAAAAACAAAAACAA 2156
DB 127035 GCCACTGCACCTCAGCTGGGGAGAGAGAGTCTCTCTCAAAAAACAAAAACAA 127093
QY 2157 AAGAATTAAGCAAAATAGACATTTGAGAGAGAACTTGAAGGGGTGAGACACGATACAG 2216
DB 127094 AGTTAAGTAGTACTTTGGGGCCCTATCAGATAGTGTCTCTCGAGCGAGGCTGGGCCCTTA 127153
QY 2217 TTTCTGTGCCACAT 2230
DB 127154 CTACAGTTTCACTT 127167

RESULT 10

US-09-345-882-1/c
; Sequence 1, Application US/09345882
; Patent No. 6359373
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; NAME/KEY: allele

```
,
, LOCATION: 89073
, OTHER INFORMATION: 5-127-261 : polymorphic base A or C
, FEATURE:
, NAME/KEY: allele
, LOCATION: 90842
, OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
, FEATURE:
, NAME/KEY: allele
, LOCATION: 93714
, OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
, FEATURE:
, NAME/KEY: allele
, LOCATION: 97122
, OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
, FEATURE:
, NAME/KEY: allele
, LOCATION: 97152
, OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
, FEATURE:
, NAME/KEY: allele
, LOCATION: 99098
, OTHER INFORMATION: 5-130-257 : polymorphic base A or G
, FEATURE:
, NAME/KEY: allele
, LOCATION: 99117
, OTHER INFORMATION: 5-130-276 : polymorphic base A or G
, FEATURE:
, NAME/KEY: allele
, LOCATION: 103806
, OTHER INFORMATION: 5-131-395 : polymorphic base A or T
, FEATURE:
, NAME/KEY: allele
, LOCATION: 106940
, OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
, FEATURE:
, NAME/KEY: allele
, LOCATION: 108106
, OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
, FEATURE:
, NAME/KEY: allele
, LOCATION: 108149
, OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
, FEATURE:
, NAME/KEY: allele
, LOCATION: 108308
, OTHER INFORMATION: 5-135-357 : polymorphic base A or G
, FEATURE:
, NAME/KEY: allele
, LOCATION: 108471
, OTHER INFORMATION: 5-136-174 : polymorphic base C or T
, FEATURE:
, NAME/KEY: allele
, LOCATION: 134134
, OTHER INFORMATION: 5-140-120 : polymorphic base C or T
, FEATURE:
, NAME/KEY: allele
, LOCATION: 134362
, OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
, FEATURE:
, NAME/KEY: allele
, LOCATION: 134374
, OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
, FEATURE:
, NAME/KEY: allele
, LOCATION: 146328
, OTHER INFORMATION: 5-143-84 : polymorphic base A or G
, FEATURE:
, NAME/KEY: allele
, LOCATION: 146345
, OTHER INFORMATION: 5-143-101 : polymorphic base A or C
, FEATURE:
, NAME/KEY: allele
, LOCATION: 150329
, OTHER INFORMATION: 5-145-24 : polymorphic base A or G
, FEATURE:
, NAME/KEY: allele
, LOCATION: 160031
, OTHER INFORMATION: 5-148-352 : polymorphic base G or T
, FEATURE:
, NAME/KEY: allele
, LOCATION: 72771..72817
, OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
, FEATURE:
, NAME/KEY: allele
, LOCATION: 72771..72817
, OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
, FEATURE:
, NAME/KEY: allele
, LOCATION: 88050..88096
, OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
, FEATURE:
, NAME/KEY: allele
, LOCATION: 88050..88096
, OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
, FEATURE:
, NAME/KEY: allele
, LOCATION: 90819..90865
, OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
, FEATURE:
, NAME/KEY: allele
, LOCATION: 90819..90865
, OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
, FEATURE:
, NAME/KEY: allele
, LOCATION: 93690..93736
, OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
, FEATURE:
, NAME/KEY: allele
, LOCATION: 93690..93736
, OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
, FEATURE:
, NAME/KEY: allele
, LOCATION: 97099..97145
, OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
, FEATURE:
, NAME/KEY: allele
, LOCATION: 97130..97177
, OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
, FEATURE:
, NAME/KEY: allele
, LOCATION: 97130..97177
, OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
, FEATURE:
, NAME/KEY: allele
, LOCATION: 99075..99121
, OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
, FEATURE:
, NAME/KEY: allele
, LOCATION: 99075..99121
, OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
, FEATURE:
, NAME/KEY: allele
, LOCATION: 99094..99140
, OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
, FEATURE:
, NAME/KEY: allele
, LOCATION: 99094..99140
, OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
, FEATURE:
, NAME/KEY: allele
, LOCATION: 103783..103828
, OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
, FEATURE:
```

FEATURE:
 NAME/KEY: allele
 LOCATION: 103783..103828
 OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
 FEATURE:
 NAME/KEY: allele
 LOCATION: 106918..106966
 OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
 FEATURE:
 NAME/KEY: allele
 LOCATION: 106918..106966
 OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
 FEATURE:
 NAME/KEY: allele
 LOCATION: 108084..108130
 OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
 FEATURE:
 NAME/KEY: allele
 LOCATION: 108127..108177
 OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
 FEATURE:
 NAME/KEY: allele
 LOCATION: 108127..108177
 OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
 FEATURE:

Query Match 7.9%; Score 308.2; DB 4; Length 162450;
 Best Local Similarity 66.7%; Pred. No. 5.6e-66;
 Matches 529; Conservative 0; Mismatches 233; Indels 31; Gaps 5;

QY 1421 AAAATAATAAAAAACACAAATTTTTTTTTTTTTTTGAGAAAGAGCTGTGCTCTTTACCC 1480
 DB AATGTAACAAGTGAACAACTATTTTTTTTGTAGGGGGAGTCTAGCTGT-TGC 86793
 QY 1481 CTGGCTGAGGGAGTGTGTGATCTCAGCTCAGTGCACACTTTCCGCTCCCGGTTCAAG 1540
 DB CGGGCTGAGTGCAGTGGCACCACCTTGGCTCAGTGCACAGCTCCGCCCTCTGGGTTCAAG 86733
 QY 1541 CAATCTCTCTCAGCTCCAGTGCAGTGGGAGTGCAGTCTCCACCATGCCCCA 1600
 DB CAATCTCTCTCAGCTCCAGTGCAGTGGGAGTGCAGTCTCCACCATGCCCCA 86673
 QY 1601 GCTGATTTTTGATTTTGTAGAGATGGGATTTCACTTTGTTGCCAAGTGGTCTCAA 1660
 DB GCTAATTTTTGATTTTGTAGAGAGAGGGGTTTCAACCATGTTAGCCAGGATAGTCTCGA 86613
 QY 1661 ACT-----TTTTGCTGTGATTAATTTGTTGTAACATTTGTTTCTTTGCTGAGGT 1708
 DB TCTCTGACCTGCTGATCCGCCACCTCTCGCTCCCAAGTGTGGGATATAGGGGTGAG 86553
 QY 1709 AGGGCCCCCAGACCAAAAAATAAATCTTAGAATCCAAATCAGTGTGTTGTTGACCA 1768
 DB CCACCCACCTGGCCAAAAAAGTAAACTATTAATAATTAAGCTTTTAAAGTGA 86493
 QY 1769 CTGTCACTTGAACACACAGTG-----TGACAGGGGCTCAGGAGTAGAGTGAATCTC 1821
 DB CATGGGAATACCAACCCCTTGAATTTCAATGATGATGATCAATTAATACTTAAATAATTTGCC 86433
 QY 1822 TGCTGAAGAGAAATAGATGAATAATTTCTCCGGGCCAGG-----CGTGGTGG 1871
 DB TGGCAGAAACAAAGCAAAATCTTTCAAAGCTTGTCTATTAATGAAGATCTGATCTTTGGCT 86373
 QY 1872 CTCATGCTGTAAATCCAGCACTTTGGAGGGCCAGAGGATGGGATCACTCAGGTCAAG 1931
 DB GGAAGTATTTGATCTCCAGCACTTTGGAGGGCCAGAGGATGGGATCACTCAGGTCAAG 86313
 QY 1932 AGTTCAAAACCAAGCTGGCCACATGGTGAACCCCGCTCTCTACTAAATAACAAAAAT 1991
 DB AGTTCAAGACCAAGCTGGCCAAATGGTGAACCCCTCGTCTCTACTAAATAAC-AAAAAT 86254

QY 1992 TAGCTAAGTGTGTGGCGCATGCTGTAAATCCAGCTACTTGGGAGGTGAGGAGGAGA 2051
 DB TAGCCAGGATGTGGCGGTGCTGTAAATCCAGCTACTTCTGGAGGTGAGGAGGAGA 86194
 QY 2052 ATTTCTTTGAACCCCGGAGGAGAGGTTGCAGTGAAGCGAGATCACACACTGCACTCCAG 2111
 DB ATTGCTCTGAACCCCGGAGGAGGAGTTGCAGTGAAGCGATATCGCACACTGCACTCCAG 86134
 QY 2112 CTTGGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAAGATTAAGCAAT 2171
 DB CCTGGCGCAGAGGTGAGACTCCGCTCTCAAAAAACAAAAACAAAAAGATGTTTGTACTTCAA 86074
 QY 2172 TAGCATTGCGAGA 2184
 DB TGGCACTTAAAGA 86061

RESULT 11
 US-09-735-934A-3
 ; Sequence 3, Application US/09735934A
 ; Patent No. 6372468
 ; GENERAL INFORMATION:
 ; APPLICANT: LI, Jiaxin et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CLO00851
 ; CURRENT APPLICATION NUMBER: US/09/735.934A
 ; CURRENT FILING DATE: 2000-12-14
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 43950
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-735-934A-3

Query Match 7.9%; Score 307.6; DB 4; Length 43950;
 Best Local Similarity 67.6%; Pred. No. 4e-66;
 Matches 491; Conservative 0; Mismatches 229; Indels 6; Gaps 4;
 QY 1442 TTTTTTTTTTTTTTTTTGAGAAAGAGTCTTGTCTCTTTTCACTCCCTGGTGGAGGCGAGTGTGT 1501
 DB TTTTTTTTTGTTGTGTGTGACAGAGTCTTGTCTCCGTCACTCCAGAGTGGAGTGCAGTGGCGT 10021
 QY 1502 GATCTCAGCTCAGTGCACACTTTGCGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAAGCTTC 1561
 DB GGTCTTGGCTCAGTGCACACTCTACCTCTCTGGGTTCAAGCAATTTCTCTGCTTCAAGCTTC 10081
 QY 1562 CCAAGTAGCTGGGACTACAGGCACTTCCACCATGCCAGCTGATTTTGTATTTTGTAGT 1621
 DB CCAATAGCTGGGATTTAGAGTGGCGCACCACTCTCTGGCTGATTTTGTATTTTGTAGT 10141
 QY 1622 AGAGATGGGATTTTCACTTTTGTGGCAAGTGTGTCTCAAACTTTTGTGCTGATAAATGT 1681
 DB --AGACGGGGTTTCAACCATTTTGGCCAGCTGGTTTCAAGTAATCCACCTCTCAGGCT 10199
 QY 1682 TGTAATTAATTTGTTTCTTTTGTGAGTAGGCGCCCGCCAGCAAAAAAATAAATTTAGA 1741
 DB CCCCAGTGTGGGATTAACAGCATGAGCCACTGGCTGGGCCCCCATGCGCTTCCATTTT 10259
 QY 1742 ATCCAAATCAGTGTGTTGTTGTTGACCACTGCTCACTTGAAGAACACAGTGTGACAGGCGCC 1801
 DB TGTATGTTGTCCTCCGCAATTAGAGCCATATTTCTTGGATGTTTCCATTGGGTATTTAGTCT 10319
 QY 1802 TCAGGATGAGGTGATCTCTGCTCGAAGAGAAATAGAAATGAAATAATTTCTCGGGGCA 1861
 DB GAGACAGCATCTTAGCTCCGTTGGGTGCCACGCTTGTACAGAAATCCCTGATTTCTGGGCA 10379
 QY 1862 GCGTGTGGCTCATGCTGTAAATCCAGCACTTTGGAGGCGCAAGGATGTTGGATCACC 1921
 DB GGCACGGTGGCTCACACCTGTAAATCCAGCACTTTGGAGGCGCAAGGCGGATCA-- 10437


```

Db 50106 -----GCATTGGTTACCTTTTATGGTTGGCACT--GGATGGATGGATGGCCATTTA 50058
Qy 1800 CCTCAGGAGTAGAGTGATCTCTGCTCGAAAGAAATAGAAATGAAATATTTCTCCGGG 1859
Db 50057 TGGTAGTGATGCAACATTTCTCTCAAAATAAAATTAAGTCAGTTTAAAGATAAGCCAGGC 49998
Qy 1860 CAGCGTGGTGGCTATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1919
Db 49997 CGGCGACGGTGGCTGACGCTGTTGATCCAGCAGTTTGGGAGGCTGAGGAGGTTGGAACA 49938
Qy 1920 CCTGAGTCAAGGAGTTCAAACACAGCTGGCCACATGCTGAAACCCCGTCTCTACTAAA 1979
Db 49937 CCTGAGTCAAGGAGTTCAGACAGGCTGACCAATGAGGAGAAACCCCGTCTCTACTAAA 49878
Qy 1980 AATACAAAAAATAGCTAAGTGGTGGCGCATGCTGTAAATCCAGCTACTTGGGAGGG 2039
Db 49877 AATAC--AAAATAGCAGGTGTGGTGGCGCATGAACTGTAATCCAGCTACTCCTGGGAAGC 49820
Qy 2040 TGAGGAGGAGAAATTTCTTGAACCCGGGAGGAGGAGGTTGCAAGTGAAGGAGATCAACC 2099
Db 49819 TGAGGAGGAGAAATCGGTTGAAACCCGAGGAGGAGGTTGCGGTGAGCAGATCGTGCC 49760
Qy 2100 ACTCACTCCAGCTGGGGA-GAGAGCGAGCTTCTCTCAAAACCAAAAAACAA 2156
Db 49759 ATTGCACTCCAGCTGGGCAACAAAGGAAACTCTCTCAAAATAAATAAATAA 49702

RESULT 14
US-09-820-924-3/c
; Sequence 3, Application US/09820924
; Patent No. 655351
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001213
; CURRENT APPLICATION NUMBER: US/09/820,924
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 39982
; TYPE: DNA
; ORGANISM: Human
US-09-820-924-3

Query Match 7.8%; Score 304; DB 4; Length 39982;
Best Local Similarity 56.3%; Pred. No. 3e-65;
Matches 829; Conservative 0; Mismatches 595; Indels 48; Gaps 12;

Qy 735 TATTTTTTTTCTGAAACAGGGTATCACTCTGTTGCCAGGCTGGAGTACAGTGGCGTAA 794
Db 32734 TATTAATTTTTTTCAGATGGAGTCTCACTTTGTGCGCCATGCTGAGTGCAGTGACGTGA 32675
Qy 795 TAATAGTCACTGAGGCTCC-CCTCTGGGCTCAAGCAATCCGCTGGCTCAGCATCCT 853
Db 32674 TCTTGGCTCACTAACAACCTCCACCTCCGGGTTCAAGCTATTTCTGTGCTCAGCCTCCC 32615
Qy 854 GAGTAGTGGGACTACAGGCTTGTGCCACAGGCCCAGCTAAGTTTAAATAATGATTTT 913
Db 32614 GAGTAGTGGGATACAGGCAACCCACACACCACTGCTGGCTAAATTTTT-----GTACTTT 32561
Qy 914 TGGTATAGAGGAGTCTTGTCTATGTTGCTCAGGCTGTATTTTTTATGTTGAGACAAGTTC 973
Db 32560 TAGTAGATGGGTTTCCACCATGTTGGCCGCTGGCTTGAATCTTGGTCTCAAGTGA 32501
Qy 974 TCATATGTTGCCATGATCCGCCACCTCCATCTCCCAAGATGCTCATC--TTATCTGTT 1031
Db 32500 TCCGCTCTTGGCTCCCAAGTGTGTTGAGTATACAGGTGTGTGCCACCAAGCCAGCT 32441
Qy 1032 CATTAGTCACTGACAGACATTTAGTTGTTTCCACTTTTGGCAATATGAAATAACT 1091

```

```

Db 32440 AATTTTGTATTTTTTAGTGAGAAAGGGTTTCAACATGTTGGCCAGGCTGCTCGAACT 32381
Qy 1092 CCAGTGAATATTCATGTATACATTTGTGTGGCATATGTTTTCATTT----- 1138
Db 32380 CTTGACCTCTGTATTCGGCGCCTCAGCCTCCCAAAGTGTGGATTAACAAGCCAGCCC 32321
Qy 1139 -----CTGTTTGGTATATATCTAGGATGGAATTCCTGGATCCCGGTAATATATTTTGACA 1193
Db 32320 TGATACAGTGGTTTAAAAATGAGGAAGATCTTCAAAACCAATGTGGAATCATTTTCTTA 32261
Qy 1194 GGCAGAGTTTCAGGGGAGAAAAAATCTTGGGAAAATGAGCATGTTTAG-----AAA 1243
Db 32260 ATAAATTTTAACTGAAGAAATGCAAAAGGAAAGCAGAGTAGTCTAGGGTACTACTTCTT 32201
Qy 1244 TCAGCAAGAGTGCAGGGTTTTTCGGAGTTTTTATTTATATTTCTGTTGACAAATGTGTCAG 1303
Db 32200 GCTGTGAGAAAGAGGGAAGATAGCATACTTGTCTGCGCTGTTTGTCTAAACGAAAC 32141
Qy 1304 TTTGATGAAGATACAAGTTTACTTAAGTGAGAGTGAAGATTAAGGCTTGAATAGGGCGT 1363
Db 32140 CAGGAGGAGGAACAG--AAGTGAACGAGACTGCTTACCTACAGGGGGTAGGGGAGA 32083
Qy 1364 TCAGAGTAAATCATGAAGCACTTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAA 1433
Db 32082 AAACAGGCAAAAGGAGAGTACTGTTATTTGTATGTTCTGACTTTTGGAGCTACATTA 32023
Qy 1424 ATAAATAAAATCACAATTTTTTTTTTTTTTTTTTTTGTAGAAAGAGTCTTCTTTTACCCCTG 1483
Db 32022 TGTCTTTTATTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 31964
Qy 1484 GCTGGAGGCGAGTGGTGTGATCTCAGCTCACTGCAACTTTTCGCTCCCTCC--GGGTTCAAGCA 1542
Db 31963 GCTGGAGTGCAGTGGTGTGATCAAGCTCACTGAGCTCGACTCCAGGGCTCAGGTG 31904
Qy 1543 ATTCTCTGCTTCCAGCTCCCAAGTAGCTGGGACTCAGGCACTTCCACCATGCCCAGC 1602
Db 31903 ATCTCCCACTCAGCTCCAGCTCCAGCTTCCAGGCTTCCAGGTGCACTACACACCCAGC 31844
Qy 1603 TGAATTTTGTATTTTGTAGTAGAGTGGGATTTTCACTTTGTTGGCCAGCTGCTCTCAAC 1662
Db 31843 TAAATTTTGTATTTCTTGTAGACA--GAGGTTTCCCATGTTGCCAGGCTGGTCTCAAAC 31785
Qy 1663 T-TTTTGTCTGTCATAATTTGTGTAACCTATTGTTCTTTTGTGAGGTAGGCGCCCCAGAC 1721
Db 31784 TCTGGGCTCAAGCAATCCATCTGCTTGGCTCCCAAAGTGTAGGATTAACAGGCATGA 31725
Qy 1722 CAAAAAATAAATCTTGAATCCAAATCAGTGTGTTGTTGACCACTGCTCACTTGAGA 1781
Db 31724 GACACTGTGCCAGCCACGTTGTTTCAACATTCAAAAACAACAGCAACAATAACAAACG 31665
Qy 1782 ACCACAGTGTGACAGGGCTCAGGAGTAGAGTGTCTCTGCTCGAAAGAGAAATAGAA 1841
Db 31664 CTCAAAAACCAAAACACGAAATCAGCAAGATGAGAGGGAATCCCTAAAGTGAATAAATA 31605
Qy 1842 TGAATAATTTCTCGGCGCCAGGGTGTGCTCATGCTGTATTCATGCCAGCACTTTGGGAG 1901
Db 31604 TGAATCCAA---CTAGGCTGGCGCGGTGCTCACACTGTATTCAGCACTTTGGGAG 31548
Qy 1902 GCCAAGCATGTGATCACTGAGGTGAGGTTTCAAAACAGCTGCGCAACATGTTGA 1961
Db 31547 GCCGAGCGGGCGGATCAC--GAGGTGAGAGATCGAGACCATCTCTGGCTTAACACCATGA 31490
Qy 1962 AACCCGCTCTACTATAAATAAATAAATAAATAGTAACTGTGGTGGCGCATGCTGTAAT 2021
Db 31489 AACCCCATCTCTATAAATAAATAAATAAATAGCCGGCGGTGGTAGCGGCACCTGTAGT 31430
Qy 2022 CCAGCTACTTGGAGGCTGAGCGAGGAGAAATTTCTTGAACCCCGGAGGAGAGGTTGCA 2081
Db 31429 CCAGCTACTTGGAGGCTGAGCGAGGAGAAATGCGGTGAACCCCGGAGGAGAGGTTGCA 31370
Qy 2082 GTGAAGCGAGATCACCACTGCACTCAGCTCGGGGAGAGGAGACTTCTCTCAA 2141
Db 31369 GTGAGCTGAGATAGTGCACTTGCACTCCGGCTGGGCGAAAGAGCAAGACTCCGCTCTCAA 31310

```

QY 2142 ARAACACAAAAACAARAGATTAGCAATTA 2173
|||||
Db 31309 AAAAAAARAAAAAAGAACCACTA 31278

RESULT 15

US-09-735-934A-3/c
; Sequence 3, Application US/09735934A
; Patent No. 6372468
; GENERAL INFORMATION:
; APPLICANT: LI, Jiayin et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000851
; CURRENT APPLICATION NUMBER: US/09/735,934A
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 43950
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-934A-3

Query Match 7.7%; Score 303; DB 4; Length 43950;
Best Local Similarity 68.2%; Pred. No. 5.6e-65;
Matches 492; Conservative 0; Mismatches 200; Indels 29; Gaps 4;

QY 1441 TTTTNTTTTTTTTTGAGAAAGAGTCTGCTCTTTTCCACCTGGCTGGAGGGCAGTGGTG 1500
|||||
Db 6167 TTTTNTTGTGTTGTTAGAGAGTTTCACTCTGTCAACCCAGGCTGGAGTACAGTAGTG 6108
QY 1501 TGATCTCAGCTCACTGCACTTTTCGCTCCGGGGTTCAGCAATTCCTCTGCTTCAAGCCT 1560
|||||
Db 6107 AGATCTGGGCTCAACCAA-GTCCGGCTCCCGGGTTCAAGCAGTTCTCTGCTCAGCCT 6049
QY 1561 CCCAAGTAGCTGGGACTACAGGCACATCCACCACATGCCAGCTGATTTTGTATTTTTTAG 1620
|||||
Db 6048 CCAGTAGCTGGGATTACAGGATGGGCCACCACACCCGACTAATTTTGTATTTTTTAG 5989
QY 1621 TAGAGATGGGATTTCACTTTTGTGGCA-AGCTGGTCTCAAACTTTTGTGTCATAATT 1679
|||||
Db 5988 TAGAGACGGAGTTTCAACCATGTTGGCCAGGATGGTCTCAAACTCCGACTCAGGTGATC 5929
QY 1680 GTTGTAACTATTGTTCTCTTTTGTAGTAGGAGCCGCCAGACCAAAAAATAAATCTTA 1739
|||
Db 5928 TGCCCGCTCGGTCTCTCAAAAGTGTGGGATTACAGCGGTGAGCCACTGCAACCCGGCCTA 5869
QY 1740 GAATCCAAATCAGTGTGTTGTTGACCACTGTCACTTGAGAACCAACAGTGTGACCAAGG 1799
|||
Db 5868 AATAGTAATTTTATGTTATGTATATTTTACCAATTAATAATTTTAAATAAATA 5809
QY 1800 CTTCAAGAGTAGAGTGATCTGTGCGAAGAGAAATAGAAATATTTCCCGGGC 1859
|||
Db 5808 CC-----ACCACCAACAACAAAAAACAATGATAGGC 5775
QY 1860 CAGCGTGTGGTCTATGCTGTAAATCCAGCACTTTGGAGGCCAAGCATGTGGATCA 1919
|||||
Db 5774 GGTGCCCGGTGGGTCTCATCTGTAAATCCAGCACTTTGGAGGCCAAGGTGGGGCGATCA 5715
QY 1920 CTTGAGGTCAAGAGTTCAAAACCAAGCTGGCCCAACATGTTGAAACCCCGTCTCTACTAAA 1979
|||||
Db 5714 CTTGAGGTCAAGAGTTGAAACCAAGCTGGCCCAACATGTTGAAACCCCGTCTCTACTAAA 5655
QY 1980 AATACAAAAAATAGTAAAGTGTGGCGCATGCTGTAAATCCCACTACTTTGGAGGG 2039
|||||
Db 5654 AATAC-AAAAATTAGCCAGGCCTGGTCCCGGTGCTGTAAATCCCACTACTCTCAGGAGGC 5596
QY 2040 TGAGGCAGGAGAAATTTTGAACCCGGAGGAGGTTTCAGTGAAGCAGATCACACC 2099
|||||
Db 5595 TGAGGCAGGAGAAATTTGTTGAACCCGGAGGAGGTTGCCGTGAGCCAAAGATCATGCC 5536

QY 2100 ACTGCACTCCAGCCTGGGGAGAGAGCGAGACTTCTCTCAAAAAACAACAAAAAG 2159
|||||
Db 5535 ACGCACTCCAGCCTGGGTGACAGAGTAAGATTCTGTCTAAAAAAGAAAAA 5476
QY 2160 A 2160
Db 5475 A 5475

Search completed: May 13, 2004, 17:18:26
Job time : 294.273 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 09:06:53 ; Search time 2243.23 Seconds
(without alignments)
7909.918 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3910
Perfect score: 3910
Sequence: 1 aggaattcatcatcttaaat.....cggaacaggtcacccgttc 3910

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2947324 seqs, 2269024515 residues
Total number of hits satisfying chosen parameters: 5894648

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3910	100.0	4150	14	US-10-016-725-15
2	3881.8	99.3	37698	13	US-10-087-192-10
3	630.6	16.1	631	13	US-10-027-632-24494
4	630.6	16.1	631	16	US-10-027-632-24494
5	460	11.8	59725	13	US-10-087-192-814
6	450.4	11.5	617	10	US-09-792-468B-1
7	407.4	10.4	9631	15	US-10-017-161-1569
8	360.4	9.2	23618	13	US-10-087-192-4
9	353.6	9.0	36296	13	US-10-240-425-1584
10	349.8	8.9	60815	13	US-10-087-192-52
11	349.8	8.9	12394	16	US-10-181-875-10
12	349.8	8.9	13216	10	US-09-764-891-7810
13	347.4	8.9	29163	10	US-09-764-891-7809
14	342.8	8.8	26928	9	US-09-880-107-2278

15	342.8	8.8	26928	13	US-10-235-192A-34	Sequence 34, Appl
16	342.8	8.8	26928	15	US-10-020-141-7	Sequence 7, Appl
17	342.8	8.8	26928	15	US-10-017-631-1	Sequence 1, Appl
18	338.6	8.7	144035	13	US-10-087-192-322	Sequence 322, App
19	337.2	8.6	3273	16	US-10-012-697-1406	GENERAL INFORMATION
20	337	8.6	109906	13	US-10-235-192A-31	Sequence 31, Appl
21	334.8	8.6	108182	13	US-10-087-192-1168	Sequence 1168, Ap
22	333.8	8.5	175077	13	US-10-087-192-1168	Sequence 1168, Ap
23	333.6	8.5	35425	15	US-10-017-161-2429	Sequence 2429, Ap
24	333.6	8.5	35425	16	US-10-292-798-2069	Sequence 2069, Ap
25	332.2	8.5	3158	13	US-10-027-632-115030	Sequence 115030, Ap
26	332.2	8.5	3158	16	US-10-027-632-115030	Sequence 115030, Ap
27	330.6	8.5	3158	13	US-10-027-632-115030	Sequence 115030, Ap
28	330.6	8.5	3158	16	US-10-027-632-115030	Sequence 115030, Ap
29	330.2	8.4	198522	13	US-10-087-192-244	Sequence 244, App
30	329.6	8.4	9372	9	US-09-764-877-3049	Sequence 3049, Ap
31	329.6	8.4	9372	9	US-09-764-877-3244	Sequence 3244, Ap
32	329.6	8.4	9372	16	US-10-242-515-3049	Sequence 3049, Ap
33	329.6	8.4	9372	16	US-10-242-515-3244	Sequence 3244, Ap
34	328.8	8.4	2130	16	US-10-108-260A-2246	Sequence 2246, Ap
35	328.8	8.4	75252	13	US-10-087-192-904	Sequence 904, App
36	328.6	8.4	35425	15	US-10-017-161-2429	Sequence 2429, Ap
37	328.6	8.4	35425	16	US-10-292-798-2069	Sequence 2069, Ap
38	328.4	8.4	96594	13	US-09-997-722-154	Sequence 154, App
39	328	8.4	91760	13	US-10-087-192-844	Sequence 844, App
40	327.6	8.4	12919	15	US-10-017-161-787	Sequence 787, App
41	327.6	8.4	17397	9	US-09-764-869-1945	Sequence 1945, Ap
42	327.6	8.4	17397	15	US-10-091-504-1945	Sequence 1945, Ap
43	327.6	8.4	17397	16	US-10-227-577-1945	Sequence 1945, Ap
44	327.6	8.4	19334	9	US-09-764-869-1943	Sequence 1943, Ap
45	327.6	8.4	19334	15	US-10-091-504-1943	Sequence 1943, Ap

ALIGNMENTS

RESULT 1

US-10-016-725-15
; Sequence 15, Application US/10016725
; Publication No. US20020151018A1
; GENERAL INFORMATION:
; APPLICANT: Prouty, Stephen
; APPLICANT: Zhang, Lin
; APPLICANT: Steen, Kurt
; TITLE OF INVENTION: Stearyl-CoA Desaturase Promoter
; FILE REFERENCE: J6J2065
; CURRENT APPLICATION NUMBER: US/10/016,725
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 4150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-016-725-15

Query Match	100.0%	Score 3910;	DB 14;	Length 4150;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3910;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	AGGAATTCATCCATTTAAATCATACATAATTTAATGCTTTTAGTATATTCACAGGTTGTC	60	
DB	1	AGGAATTCATCCATTTAAATCATACATAATTTAATGCTTTTAGTATATTCACAGGTTGTC	60	
QY	61	ATCCATCAATCCATTTTAGAACAGTTTTACTCCAAAATAAACCCCTGCTTCCTT	120	
DB	61	ATCCATCAATCCATTTTAGAACAGTTTTACTCCAAAATAAACCCCTGCTTCCTT	120	
QY	121	AGCCATACCCCCCAACATCTCCATCTCTTCCAGCCCTGGGCAACCCACCAATCTAC	180	
DB	121	AGCCATACCCCCCAACATCTCCATCTCTTCCAGCCCTGGGCAACCCACCAATCTAC	180	
QY	181	TTTCTGCTCTATAAATTTGCCAATTCGGACATTTCTGACATTTATATAATGGAGCAACAT	240	

Db ||||| 181 TTCTGTCTCTATAAATTTGCCAATTCGGCAATTCATATAATGGAGCAACACAT 240
Qy ||||| 241 GTGAGCTTTGTGACTGGCTGCTTTCACCTAGCAATCTATTTTAAAGGCTCAATATGTTA 300
Db ||||| 241 GTGAGCTTTGTGACTGGCTGCTTTCACCTAGCAATCTATTTTAAAGGCTCAATATGTTA 300
Qy ||||| 301 CAGTACTTAGCAGTACTCAATCTTTTATTTTATTTCTCAAAATGGFATTCOACTGHTGGGTAT 360
Db ||||| 301 CAGTACTTAGCAGTACTCAATCTTTTATTTTATTTCTCAAAATGGFATTTCCACTGTGTGGGTAT 360
Qy ||||| 361 CCCATATCATATATTAGAGACAGGTCTCTCACTCTGTCAACCAGGCTGGAGTGCAGTGC 420
Db ||||| 361 CCCATATCATATATTAGAGACAGGTCTCACTCTGTCAACCAGGCTGGAGTGCAGTGC 420
Qy ||||| 421 ACAATCATAGCTCACTGTAACTCAAACTCCTGGGCTCAAGTGATCTACTACCTCAGCC 480
Db ||||| 421 ACAATCATAGCTCACTGTAACTCAAACTCCTGGGCTCAAGTGATCTACTACCTCAGCC 480
Qy ||||| 481 TCCAGAGTACTAGGACTACAGGCACACACAGCCATACCTGGCTAATTTTTTTTAAAT 540
Db ||||| 481 TCCAGAGTACTAGGACTACAGGCACACACAGCCATACCTGGCTAATTTTTTTTAAAT 540
Qy ||||| 541 TTTCATTTTATGATTTCAATTTCTTTCTTTTGTGTTGTTGTTGTTGTTGATAGGGTCTC 600
Db ||||| 541 TTTCATTTTATGATTTCAATTTCTTTCTTTTGTGTTGTTGTTGTTGTTGATAGGGTCTC 600
Qy ||||| 601 ACTTTGTTTACCAGGCTGGAGGCACTGTCATGGTGCACAGCTGAGAGCCTTGACTTCT 660
Db ||||| 601 ACTTTGTTTACCAGGCTGGAGGCACTGTCATGGTGCACAGCTGAGAGCCTTGACTTCT 660
Qy ||||| 661 GGCTCAAGTGATCT 720
Db ||||| 661 GGCTCAAGTGATCT 720
Qy ||||| 721 CCATGCTCGGCTGATATTTTCTTTGAAACAGGGTATCACTCTGTTGCCAGGCTGGA 780
Db ||||| 721 CCATGCTCGGCTGATATTTTCTTTGAAACAGGGTATCACTCTGTTGCCAGGCTGGA 780
Qy ||||| 781 GTACAGTGGCGTAATAATAGCTCACTGCAAGCTCCCTCTGGGCTCAAGCAATCGCTG 840
Db ||||| 781 GTACAGTGGCGTAATAATAGCTCACTGCAAGCTCCCTCTGGGCTCAAGCAATCGCTG 840
Qy ||||| 841 GCCTCAGCATCTGTAGTCTGGACTACAGGCTTGTGCCACAGGCCCCAGCTAAGTTTT 900
Db ||||| 841 GCCTCAGCATCTGTAGTCTGGACTACAGGCTTGTGCCACAGGCCCCAGCTAAGTTTT 900
Qy ||||| 901 AAAAAATGATTTTGGTATAGAGGAGTCTTGCTATGTTGCTCAGGCTGTATTTTATG 960
Db ||||| 901 AAAAAATGATTTTGGTATAGAGGAGTCTTGCTATGTTGCTCAGGCTGTATTTTATG 960
Qy ||||| 961 TTGAGCAAGGCTCTACTATGTTGCCATGATCCCCCAGCTCCACTCCCAAGTCTCA 1020
Db ||||| 961 TTGAGCAAGGCTCTACTATGTTGCCATGATCCCCCAGCTCCACTCCCAAGTCTCA 1020
Qy ||||| 1021 TCTTATCTGTCTATTTAGTCTGATGACAGATTTAGTGTGTTTCCACTTTTGAACATTA 1080
Db ||||| 1021 TCTTATCTGTCTATTTAGTCTGATGACAGATTTAGTGTGTTTCCACTTTTGAACATTA 1080
Qy ||||| 1081 TGAATAATCTCCAGTGAAATTTCAATGATPACATTTGTGTGGCAATGATTTTCAATTTCT 1140
Db ||||| 1081 TGAATAATCTCCAGTGAAATTTCAATGATPACATTTGTGTGGCAATGATTTTCAATTTCT 1140
Qy ||||| 1141 GTTGGGTTTATATCTAGGAGTGAATTTCTGGATCCGGGTATATTTTACAGGCGAGAG 1200
Db ||||| 1141 GTTGGGTTTATATCTAGGAGTGAATTTCTGGATCCGGGTATATTTTACAGGCGAGAG 1200
Qy ||||| 1201 TTCAGGGGAGAAAACTTCGGAAAAATGAAGCATGTTTGAATAATCAGCAAGAGTGCAGG 1260
Db ||||| 1201 TTCAGGGGAGAAAACTTCGGAAAAATGAAGCATGTTTGAATAATCAGCAAGAGTGCAGG 1260
Qy ||||| 1261 GTTTTTCGGAGTTTTTATTTTATATTTCTGTTTGACAAATGTGAGTTTGATGAAGATACAAG 1320

Db ||||| 1261 GTTTTTCGGAGTTTTTATTTTATATTTCTGTTTGACAAAATGTCAGTTTGTATGAAGATACAAG 1320
Qy ||||| 1321 TTATACTAAGTGAAGATGAAGTGAANAATTAAGGCTGGAAATAGGGGCTTCAGAGTAAAATCATGA 1380
Db ||||| 1321 TTATACTAAGTGAAGATGAAGTGAANAATTAAGGCTGGAAATAGGGGCTTCAGAGTAAAATCATGA 1380
Qy ||||| 1381 AGCACTTTTGAATACCAAAATTAAGGAGCTTCGCTGTAAACAAAAATAATAAAAAATCAAA 1440
Db ||||| 1381 AGCACTTTTGAATACCAAAATTAAGGAGCTTCGCTGTAAACAAAAATAATAAAAAATCAAA 1440
Qy ||||| 1441 TTTTTTTTTTTTTTTTTGTGAGAAAGAGTCTTGCTCTTTTCACTCCCTGGCTGAGGGCAGTGTG 1500
Db ||||| 1441 TTTTTTTTTTTTTTTTTGTGAGAAAGAGTCTTGCTCTTTTCACTCCCTGGCTGAGGGCAGTGTG 1500
Qy ||||| 1501 TGATCTCAGCTCACTGCNAACCTTCGCTCCCGGTTCAAGCAATTTCTCCTGCTTCAGGCT 1560
Db ||||| 1501 TGATCTCAGCTCACTGCNAACCTTCGCTCCCGGTTCAAGCAATTTCTCCTGCTTCAGGCT 1560
Qy ||||| 1561 CCCAAGTAGCTGGGACTACAGGCACCTTCCCACCATGCCCCAGCTGATTTTGTATTTTAG 1620
Db ||||| 1561 CCCAAGTAGCTGGGACTACAGGCACCTTCCCACCATGCCCCAGCTGATTTTGTATTTTAG 1620
Qy ||||| 1621 TAGAGATGGGATTTTCACTTTGTTGGCCTAAGCTGCTCAAACTTTTCTGCTGTCAATTTG 1680
Db ||||| 1621 TAGAGATGGGATTTTCACTTTGTTGGCCTAAGCTGCTCAAACTTTTCTGCTGTCAATTTG 1680
Qy ||||| 1681 TTGTAACTATTTGTTCTTTTCTGCTGAGGTAGGGCCCCCAGACCAAAAAATAAATCTTTAG 1740
Db ||||| 1681 TTGTAACTATTTGTTCTTTTCTGCTGAGGTAGGGCCCCCAGACCAAAAAATAAATCTTTAG 1740
Qy ||||| 1741 AATCCAAATCAGTGTGTTGGTTTGACCACTGTCTCTTGAGAACCAACAGCTGCCAGGGC 1800
Db ||||| 1741 AATCCAAATCAGTGTGTTGGTTTGACCACTGTCTCTTGAGAACCAACAGCTGCCAGGGC 1800
Qy ||||| 1801 CTCAGGAGTGAAGTGTATCTCTGCTCGAAAGAGAAATAGAAATTAATTTCTCCGGGCC 1860
Db ||||| 1801 CTCAGGAGTGAAGTGTATCTCTGCTCGAAAGAGAAATAGAAATTAATTTCTCCGGGCC 1860
Qy ||||| 1861 AGGCGTGTGGCTCATGCTGTAAATCCAGACATTTTGGAGGCCAAGGATGTGGATCAC 1920
Db ||||| 1861 AGGCGTGTGGCTCATGCTGTAAATCCAGACATTTTGGAGGCCAAGGATGTGGATCAC 1920
Qy ||||| 1921 CTGAGCTCAGGAGTTCAAAAACAGCCTCGCCCAATAGTGTGAAACCCCTCTCTATAAAA 1980
Db ||||| 1921 CTGAGCTCAGGAGTTCAAAAACAGCCTCGCCCAATAGTGTGAAACCCCTCTCTATAAAA 1980
Qy ||||| 1981 ATACAAAAATTTAGCTAAGTGTGTGGCATGCTCTGTAAATCCAGCTACTTTGGAGGGT 2040
Db ||||| 1981 ATACAAAAATTTAGCTAAGTGTGTGGCATGCTCTGTAAATCCAGCTACTTTGGAGGGT 2040
Qy ||||| 2041 GAGGCGAGAGAAATTTCTTGAACCCCGGAGGCGAGGTTGCACTGAAGGAGATCACCA 2100
Db ||||| 2041 GAGGCGAGAGAAATTTCTTGAACCCCGGAGGCGAGGTTGCACTGAAGGAGATCACCA 2100
Qy ||||| 2101 CTGCACTCCAGCCTCGGCGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAA 2160
Db ||||| 2101 CTGCACTCCAGCCTCGGCGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAA 2160
Qy ||||| 2161 ATTAAGCAAAATTTAGCAATTTGACAGAGAACCTGAGGGGCTCAGACCAGTACAGATTTC 2220
Db ||||| 2161 ATTAAGCAAAATTTAGCAATTTGACAGAGAACCTGAGGGGCTCAGACCAGTACAGATTTC 2220
Qy ||||| 2221 TGTGCCCATATGCAAAAGTACTTCTGAGGCGATGACTGGATGAGCTGCCATCTGAAATCA 2280
Db ||||| 2221 TGTGCCCATATGCAAAAGTACTTCTGAGGCGATGACTGGATGAGCTGCCATCTGAAATCA 2280
Qy ||||| 2281 TCCAGTCTTTGTTCAGAACTTTTCAACCGGACAGGGAGCCAGACTTGAATGCACTCTCT 2340
Db ||||| 2281 TCCAGTCTTTGTTCAGAACTTTTCAACCGGACAGGGAGCCAGACTTGAATGCACTCTCT 2340
Qy ||||| 2341 GGTCTAGTGCAGAGAGTTGGCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTT 2400
Db ||||| 2341 GGTCTAGTGCAGAGAGTTGGCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTT 2400

183 TCCTGCTCTATAAATTTGCCAATTCCTGGACATTTTCATATAAATGGAAGCAAAACAATGT 242
 Db
 5460 TCCTGCTCTATAAATTTGCCAATTCCTGGACATTTTCATATAAATGGAAGCAAAACAATGT 6519
 QY
 243 GAGACTTTGTGACTGGCTGCTTTTCACTAGACATTCATATTTTAAAGGCTCATATATGTACA 302
 Db
 6520 GAGACTTTGTGACTGGCTGCTTTTCACTAGACATTCATATTTTAAAGGCTCATATATGTACA 6579
 QY
 303 GTACTTAGCAGTACTTCATTTCTTTTATTTCTCAATGGTATTCCTCACTGTGGGTATCC 362
 Db
 6580 GTACTTAGCAGTACTTCATTTCTTTTATTTCTCAATGGTATTCCTCACTGTGGGTATCC 6639
 QY
 363 CATATCATATTTATTAGACAGGTTCTCACTGTCTCACCCAGGCTGGAGTGCAGTGGCAC 422
 Db
 6640 CATATCATATTTATTAGACAGGTTCTCACTGTCTCACCCAGGCTGGAGTGCAGTGGCAC 6699
 QY
 423 AATCATAGCTCACTGTAACTCAAACTCTCTGGGCTCAAGTGATCTTACTACCTCAGCCTC 482
 Db
 6700 AATCATAGCTCACTGTAACTCAAACTCTCTGGGCTCAAGTGATCTTACTACCTCAGCCTC 6759
 QY
 483 CAGAGTAGCTAGACTACAGGCACACACAGCCATACCTGGCTAATTTTTTTTTTAAATTT 542
 Db
 6760 CAGAGTAGCTAGACTACAGGCACACACAGCCATACCTGGCTAATTTTTTTTTTAAATTT 6819
 QY
 543 TCATTTTATGTAATTCATTTCTTTTCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 602
 Db
 6820 TCATTTTATGTAATTCATTTCTTTTCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6879
 QY
 603 TTTGTTTACCAGGCTGAGGGCAGTGGGATGGTGACAGCTGAGCAGCCTTGACCTTCCTGG 662
 Db
 6880 TTTGTTTACCAGGCTGAGGGCAGTGGGATGGTGACAGCTGAGCAGCCTTGACCTTCCTGG 6939
 QY
 663 GCTCAAGTGTATCTCTGCTCCCTCAGCTCCCAAGTAGCTGGGACTPACAAACACAGTGCACC 722
 Db
 6940 GCTCAAGTGTATCTCTGCTCCCTCAGCTCCCAAGTAGCTGGGACTPACAAACACAGTGCACC 6999
 QY
 723 ATGCTCTGGCTGATATTTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 782
 Db
 7000 ATGCTCTGGCTGATATTTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 7059
 QY
 783 ACAGTGGGTAATTAATAGCTCACTGCGAGCTCCCTCTCTGGGCTCAAGCAATCCGCTGGC 842
 Db
 7060 ACAGTGGGTAATTAATAGCTCACTGCGAGCTCCCTCTCTGGGCTCAAGCAATCCGCTGGC 7119
 QY
 843 CTACAGCATCTGAGTAGCTGGGACTACAGGCTGTGGCCACAGGCCCCAGCTAAAGTTTAA 902
 Db
 7120 CTACAGCATCTGAGTAGCTGGGACTACAGGCTGTGGCCACAGGCCCCAGCTAAAGTTTAA 7179
 QY
 903 AAAATGATTTTGGTATAGAGGAGTCTTGCTATGTTGCTCAGGCTGTAATTTTATTTGTT 962
 Db
 7180 AAAATGATTTTGGTATAGAGGAGTCTTGCTATGTTGCTCAGGCTGTAATTTTATTTGTT 7239
 QY
 963 GAGACAGGCTCTCACTATGTTGCCATGATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTC 1022
 Db
 7240 GAGACAGGCTCTCACTATGTTGCCATGATCCCTCCCTCCCTCCCTCCCTCCCTCCCTC 7299
 QY
 1023 TTATCTGTTTCAATAGTACAGTGTGACAGATTTAGGTTGTTTCCACTTTTGTGACCAATTATG 1082
 Db
 7300 TTATCTGTTTCAATAGTACAGTGTGACAGATTTAGGTTGTTTCCACTTTTGTGACCAATTATG 7359
 QY
 1083 AATAATCTCCAGTGAATTAATGATATACATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1142
 Db
 7360 AATAATCTCCAGTGAATTAATGATATACATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 7419
 QY
 1143 TGGGTTTATATCTAGGAGTGAATGCTGGATCCCGGGTATATTTTGCACAGGAGAGTT 1202
 Db
 7420 TGGGTTTATATCTAGGAGTGAATGCTGGATCCCGGGTATATTTTGCACAGGAGAGTT 7479
 QY
 1203 CAGGGGAAGAAAACCTTGGGAAAATGAAGCATGTTTGAAGAAATCAGCAAGAGTGCAGGGGT 1262
 Db
 7480 CAGGGGAAGAAAACCTTGGGAAAATGAAGCATGTTTGAAGAAATCAGCAAGAGTGCAGGGGT 7539
 QY
 1263 TTTTCGGAGTTTATTTATTTATTTCTGTTTGA CAATATGTGCAGTTTGTGATGAAGATCAAGTT 1322

7540 TTTTCGGAGTTTATTTTATTTATTTCTGTTGACAAATGTGCAGTTTGTATGAGATACAGTT 7599
 QY
 1323 ATACTAAGTGAAGCTGAGAAATTAAGGCTGGATAGGGGTTTCAGAGTAAATCATGAAG 1382
 Db
 7600 ATACTAAGTGAAGCTGAGAAATTAAGGCTGGATAGGGGTTTCAGAGTAAATCATGAAG 7659
 QY
 1383 CACTTTGAATACCAAAATTAAGGAGCTTGCTGTAAACAAAATAATAAAAAATCACAAT 1442
 Db
 7660 CACTTTGAATACCAAAATTAAGGAGCTTGCTGTAAACAAAATAATAAAAAATCACAAT 7719
 QY
 1443 TTTTCTTTTCTTTTGAAGAGTCTTGCTCTTTTCACTCTGCTGAGGAGGAGTGTGTG 1502
 Db
 7720 TTTTCTTTTCTTTTGAAGAGTCTTGCTCTTTTCACTCTGCTGAGGAGGAGTGTGTG 7779
 QY
 1503 ATCTCAGCTCACTGCAACTTTTCCCTCCCGGTTCAAGCAATTTCTCTGCTCAGGCTCC 1562
 Db
 7780 ATCTCAGCTCACTGCAACTTTTCCCTCCCGGTTCAAGCAATTTCTCTGCTCAGGCTCC 7839
 QY
 1563 CAAGTAGCTGGGACTACAGSCACTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1622
 Db
 7840 CAAGTAGCTGGGACTACAGSCACTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7899
 QY
 1623 GAGATGGATTTCACTTTGTTGGCAGAGTGTCTCAACTTTTCTGCTGCTGCTGCTGCTGCTGCT 1682
 Db
 7900 GAGATGGATTTCACTTTGTTGGCAGAGTGTCTCAACTTTTCTGCTGCTGCTGCTGCTGCTGCT 7959
 QY
 1683 GTAACCTATTTCTCTTTTCTGAGGTAGGCTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1742
 Db
 7960 GTAACCTATTTCTCTTTTCTGAGGTAGGCTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8019
 QY
 1743 TCCAAATCAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1802
 Db
 8020 TCCAAATCAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 8079
 QY
 1803 CAGAGTAGAGGTGATCTCTGCTGGAAGAGAAATAGAAATGAAATTTCTCGGGCCAG 1862
 Db
 8080 CAGAGTAGAGGTGATCTCTGCTGGAAGAGAAATAGAAATGAAATTTCTCGGGCCAG 8139
 QY
 1863 GCGTGTGGCTCATGCTGTAATCCAGCACTTTGGAGGCCCAAGGCATGTGATCACT 1922
 Db
 8140 GCGTGTGGCTCATGCTGTAATCCAGCACTTTGGAGGCCCAAGGCATGTGATCACT 8199
 QY
 1923 GAGGTCAAGGATTTCAAAACCAAGCTGCTCAACATGTTGTAACCCGCTCTCTACTATAAAT 1982
 Db
 8200 GAGGTCAAGGATTTCAAAACCAAGCTGCTCAACATGTTGTAACCCGCTCTCTACTATAAAT 8259
 QY
 1983 AAAAAAATTAGCTTAAGTGTGGGCTGCTGCTGTAATCCAGCTACTTTGGGAGGGTGA 2042
 Db
 8260 AAAAAAATTAGCTTAAGTGTGGGCTGCTGCTGTAATCCAGCTACTTTGGGAGGGTGA 8319
 QY
 2043 GGCAGGAGAAATTTCTGAAACCCGGGAGGAGAGGTTGCTGAGTGAAGCAGATCACACCT 2102
 Db
 8320 GGCAGGAGAAATTTCTGAAACCCGGGAGGAGAGGTTGCTGAGTGAAGCAGATCACACCT 8379
 QY
 2103 GCACTCCAGCTCGGGGAGAGGAGGACTTCTCTCAAAAAACAAAAACAAAAAGAAAT 2162
 Db
 8380 GCACTCCAGCTCGGGGAGAGGAGGACTTCTCTCAAAAAACAAAAACAAAAAGAAAT 8439
 QY
 2163 TAAGCAAAATTAGACATTTGACAGAGAACTTGAAGGGGCTCAGACCTACAGATTTCTG 2222
 Db
 8440 TAAGCAAAATTAGACATTTGACAGAGAACTTGAAGGGGCTCAGACCTACAGATTTCTG 8499
 QY
 2223 TGCACATGCCAAGTACTTCTGAGGCACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2282
 Db
 8500 TGCACATGCCAAGTACTTCTGAGGCACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8559
 QY
 2283 CAGTCTTTTTCAGAACTTTTCAACCCGACAGGGAGCCAGGACTGGAATGAGTCTCTCTG 2342
 Db
 8560 CAGTCTTTTTCAGAACTTTTCAACCCGACAGGGAGCCAGGACTGGAATGAGTCTCTCTG 8619
 QY
 2343 TCAGTCCAGAGAGTTGGCTTGAACCTCAGACCTGAGACCTGAGGCAACAAAGAGGCTGCTTAG 2402


```

Db 631 AAGTCTCCTATGTTGGCATGATCCCGCCACCTCCACTTCCCAAGTGCTCACTTATC 572
Qy 1028 TGTTCATTAGTCAGTTGACAGACATTTAGTGTGTTTCCACTTTTTCACCATTAATGAAPAA 1087
Db 571 TGTTCATTAGTCAGTTGACAGACATTTAGTGTGTTTCCACTTTTTCACCATTAATGAAPAA 512
Qy 1088 TACTCCAGTGAATATCATGTATATCATTTGTGTGGGCATATGTTTTCATTCTGTGGGT 1147
Db 511 TACTCCAGTGAATATCATGTATATCATTTGTGTGGGCATATGTTTTCATTCTGTGGGT 452
Qy 1148 TTATATCTAGGAGTGAATGCTGATCCCGGTAAATATTTTGACAGGCAGAGTTTCAGGG 1207
Db 451 TTATATCTAGGAGTGAATGCTGATCCCGGTAAATATTTTGACAGGCAGAGTTTCAGGG 392
Qy 1208 GAAGAAAAAATTTGGGAAATGAAGCATGTTTGAAGAAATCAGCAAGAGTGCAGGGGTTTTTC 1267
Db 391 GAAGAAAAAATTTGGGAAATGAAGCATGTTTGAAGAAATCAGCAAGAGTGCAGGGGTTTTTC 332
Qy 1268 GGAGTTTATTTTATATCTGTTGACAAATGTCAGTTTGATGAAGATACAAAGTTTACT 1327
Db 331 GGAGTTTATTTTATATCTGTTGACAAATGTCAGTTTGATGAAGATACAAAGTTTACT 272
Qy 1328 AAGTCAGAGTCAAGAAATTAAGGCTCGAATAGGGCGTTTCAGAGTAAATCATCAAGCATTT 1387
Db 271 AAGTCAGAGTCAAGAAATTAAGGCTCGAATAGGGCGTTTCAGAGTAAATCATCAAGCATTT 212
Qy 1388 TGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATAAAAATCAAAATTTTTTT 1447
Db 211 TGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATAAAAATCAAAATTTTTTT 152
Qy 1448 TTTTATTTTGGAGAAAGCTTCTCTCTTTCACCTGCTGGAGGCGAGTGTGATCTC 1507
Db 151 TTTTATTTTGGAGAAAGCTTCTCTCTTTCACCTGCTGGAGGCGAGTGTGATCTC 92
Qy 1508 AGCTCACTCAACTTTCGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGCCTCCCAAGT 1567
Db 91 AGCTCACTCAACTTTCGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGCCTCCCAAGT 32
Qy 1568 AGCTGGGACTACAGGCACCTTCCCAACATGCC 1598
Db 31 AGCTGGGACTACAGGCACCTTCCCAACATGCC 1

```

```

RESULT 4
US-10-027-632-24494/c
; Sequence 24494, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24494
; LENGTH: 631
; TYPE: DNA

```

```

; ORGANISM: Human
US-10-027-632-24494
Query Match 16.1%; Score 630.6; DB 16; Length 631;
Best Local Similarity 99.8%; Pred. No. 2.1e-158;
Matches 630; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 968 AAGTCTCCTATGTTGGCATGATCCCGCCACCTCCACTTCCCAAGTGCTCACTTATC 1027
Db 631 AAGTCTCCTATGTTGGCATGATCCCGCCACCTCCACTTCCCAAGTGCTCACTTATC 572
Qy 1028 TGTTCATTAGTCAGTTGACAGACATTTAGTGTGTTTCCACTTTTTCACCATTAATGAAPAA 1087
Db 571 TGTTCATTAGTCAGTTGACAGACATTTAGTGTGTTTCCACTTTTTCACCATTAATGAAPAA 512
Qy 1088 TACTCCAGTGAATATCATGTATATCATTTGTGTGGGCATATGTTTTCATTCTGTGGGT 1147
Db 511 TACTCCAGTGAATATCATGTATATCATTTGTGTGGGCATATGTTTTCATTCTGTGGGT 452
Qy 1148 TTATATCTAGGAGTGAATGCTGATCCCGGTAAATATTTTGACAGGCAGAGTTTCAGGG 1207
Db 451 TTATATCTAGGAGTGAATGCTGATCCCGGTAAATATTTTGACAGGCAGAGTTTCAGGG 392
Qy 1208 GAAGAAAAAATTTGGGAAATGAAGCATGTTTGAAGAAATCAGCAAGAGTGCAGGGGTTTTTC 1267
Db 391 GAAGAAAAAATTTGGGAAATGAAGCATGTTTGAAGAAATCAGCAAGAGTGCAGGGGTTTTTC 332
Qy 1268 GGAGTTTATTTTATATCTGTTGACAAATGTCAGTTTGATGAAGATACAAAGTTTACT 1327
Db 331 GGAGTTTATTTTATATCTGTTGACAAATGTCAGTTTGATGAAGATACAAAGTTTACT 272
Qy 1328 AAGTCAGAGTCAAGAAATTAAGGCTCGAATAGGGCGTTTCAGAGTAAATCATCAAGCATTT 1387
Db 271 AAGTCAGAGTCAAGAAATTAAGGCTCGAATAGGGCGTTTCAGAGTAAATCATCAAGCATTT 212
Qy 1388 TGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATAAAAATCAAAATTTTTTT 1447
Db 211 TGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATAAAAATCAAAATTTTTTT 152
Qy 1448 TTTTATTTTGGAGAAAGCTTCTCTCTTTCACCTGCTGGAGGCGAGTGTGATCTC 1507
Db 151 TTTTATTTTGGAGAAAGCTTCTCTCTTTCACCTGCTGGAGGCGAGTGTGATCTC 92
Qy 1508 AGCTCACTCAACTTTCGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGCCTCCCAAGT 1567
Db 91 AGCTCACTCAACTTTCGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGCCTCCCAAGT 32
Qy 1568 AGCTGGGACTACAGGCACCTTCCCAACATGCC 1598
Db 31 AGCTGGGACTACAGGCACCTTCCCAACATGCC 1

```

```

RESULT 5
US-10-087-192-814/c
; Sequence 814, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 814
; LENGTH: 59725
; TYPE: DNA

```



```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(59725)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-814

Query Match      11.8%; Score 460; DB 13; Length 59725;
Best Local Similarity 58.9%; Pred. No. 3.1e-111; Indels 111; Gaps 14;
Matches 1107; Conservative 0; Mismatches 660;

QY 365 TATCATATTATTAGACAGAGGTTCTCACTCTCTCACTCCAGGCTGGAGTGCAGTGCACAA 424
Db 33012 TTTTCTTTCTTTTGAGACAGGCTTGTCTGTCTCACTCCAGGCTGGAACCCAGTGGCACAA 32953

QY 425 TCATAGCTCACTGTAACTCAAACTCTGGCTCAAGTATCTTACTACTAGCTCAGCTCCA 484
Db 32952 TCTCAGCTCACTCAAACTCTGCTCTGCTGCTTCAAGCAATTTCTCTGCTCAGGCTCTC 32893

QY 485 GAGTAGTAGGACTACAGGCACACACAGCCATACCTGGCTAAATTTTTTTTTTAATTTTC 544
Db 32892 GAGTAGTAGGACTACAGGCCTGTGCGCACACACACAGATTAATTTTTTATATTTTATA 32833

QY 545 ATTATTATGPAATTCATTTTCT-----TTCTTTTGTGTTGTGTTTGTGATAGG 595
Db 32832 GAGATGGGGTTTCACTATGTTGGCCAGGCTGGCTTTTCTTTGTTTGTGTTTGTGACAGG 32773

QY 596 GTCCTACATTTGTTACCCAGGCTGGAGGCGAGTGGCATGGTGACAGTGA--GCAGCCTTG 653
Db 32772 GTCTTACTCTCTTTTCCAGGCTAGAGTGAGTGGTGGTCAATCATGCTCACTGCAGCTCA 32713

QY 654 ACTTCTCTGGGCTCAAGTAGTATCTCTGCTCCTCAGCTTCCCAAGTAGTGGGACTACAAACA 713
Db 32712 ACTCTTAGGCTCAGGCATCTCCACCTCGGCTTCCCAAGTGTGGAATTCACAGACA 32653

QY 714 CGTGTCACCAATG--CTGGCTGATAATTTTTTTTCTTGAACAGGGTATCATCTGTGTC 770
Db 32652 TGAGCTACCAATCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGAGACAGAGTCTTGTCTGTG 32593

QY 771 CCAGCTGGAGTACAGTGGCGTAATAATAGTCACTGACGCTCCCTCTGGCTCAAG 830
Db 32592 CCAGCTGGAGTACAGTGGCGTAATAATAGTCACTGACGCTCCCTCTGGCTCAAG 32533

QY 831 CAATCCGCTGGGCTCAGCATCTGAGTAGTGGGACTACAGGCTTGTGCCACAGGCCCA 890
Db 32532 CAATCTCTGCTCAGCATCTGAGTAGTGGGACTACAGGCTTGTGCCACAGGCCCA 32473

QY 891 GCTAAGTTTAAATAATGATTTTGGTATAGAGGAGTCTTGTCTATGTTGCTCAGGCTGT 950
Db 32472 GCTAATTTTT-----GTATTTTATAGATAGAGTGGGTTTTCAGCATGTTGGCCAGACTGG 32419

QY 951 ATTTTATTTGTTGAGACAGGCTCCTACTATGTTGCCATGATCCCTCCCTCACTCTCC 1010
Db 32418 TCTCAAACTCTTGA-----CCTCAAAATGATCGCCGCTCGGCTCTCC 32375

QY 1011 AAAGTGCTCATCTTATCTCTTCAATAGTCACTGAGTGCAGACATTTAGTGTGTTTCACTTT 1070
Db 32374 AAAGTGCTGGAATACAGCATAGGCCACTGCACCCGGGCCATTTCTTATTTTCACTTA 32315

QY 1071 TTGACCATPATGAATAATACTCCAGTGAATATTCATGTATACATTTGTGTGGCATATGT 1130
Db 32314 CAGATGAGGGAACCAAGACCTGAGAGTGAAGTCAATTTGCTTTGATGACACAGGAGT 32255

QY 1131 TTTTCTTTCTGTTGGTTTATATCTAGGAGTGAATTTGCTGATCCCGGGTAAATTTTG 1190
Db 32254 CAGCAGGGGCTGGGATTCGTGCCAGGTAAGGCTGTCTCTCAATTTTCTTTTCTTT 32195

QY 1191 ACAGGCAGAGTTTTCAGGGGAAGAAACTTGGGAAATGAAGCATGTTTGAATAATCAGCAA 1250
Db 32194 TCACITTTTTTTTTTTTTCTGAGATGGAGTTTCACTCTT---CTGCCAGGCTGGAGTGC 32138

QY 1251 GAGTGACGGGTTTTTCGGAGTTTTTATTTTATATTTCTGTTGACAAATGTGCAGTTTGATG 1310
Db 1191 GAGTGACGGGTTTTTCGGAGTTTTTATTTTATATTTCTGTTGACAAATGTGCAGTTTGATG 1310
```

RESULT 6

US-09-792-468B-1

; Sequence 1, Application US/09792468B

; Publication No. US20030157552A1

; GENERAL INFORMATION:

; APPLICANT: Hayden, Michael R.

QY	1860	CAGGCGTGGTGGCTCATGCTGTATATCCAGACACTTTGGAGGGCCAAAGCATGTGGATCA	1919
Db	10583	CAGGCATGTGGCTCATGCTGTATATCCAGACACTTTGGAGGGTGTAGGCGAGGATCA	10524
QY	1920	CCTGAGGTCAAGAGTTCAAAAACAGCTGGCCAAACATGGTGAACCCCGTCTCTACTAAA	1979
Db	10523	CCTGAGGTCAAGAGTTTGAGATCAGCCTGGCCAAACATGGTGAACCCCTGCTCTACTAAA	10464
QY	1980	AATACAAAAAATTAGCTAAGTGTGGTGGCCATGCTGTATATCCAGACTACTTTGGGAGGG	2039
Db	10463	AATAC-AAAAATTAACCGGCGTGGTGGCAGCGGCTGTATATCCAGGTACTTCGGGAGGC	10405
QY	2040	TGAGGCAGGAGAAATTTCTTAAACCCGGGAGGAGAGGTTGCAGTCAAGCGAGATCAACAC	2099
Db	10404	TGAGGCAGGAGATCGCTTGTACTTGGAAAGCAGAGGTTGCAGTGAAGCGACACTGAACC	10345
QY	2100	ACTGCACCTCAGCCTCGGGGAGAGAGCGAGACTTCTCTCAAAAAAACAACAAACAAAG	2159
Db	10344	ATTGCATCTCAGCGCTGGGTGACAGAGCAAGACTCTGTGCCAAAAAATAAATTATA	10285
QY	2160	AATTAA 2165	
Db	10284	ATTTTA 10279	
RESULT 9			
US-10-240-425-1584/C			
; Sequence 1584, Application US/10240425			
; Publication No. US20040033502A1			
; GENERAL INFORMATION:			
; APPLICANT: Williams, Amanda			
; APPLICANT: Boland, Joseph F.			
; APPLICANT: Lord, Reginald V.			
; APPLICANT: Alvarez, Chris			
; APPLICANT: Wetzel, Jon C.			
; APPLICANT: Scherf, Uwe			
; APPLICANT: Vockley, Joseph G.			
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue			
; FILE REFERENCE: 44921-5026			
; CURRENT APPLICATION NUMBER: US/10/240,425			
; CURRENT FILING DATE: 2002-09-30			
; PRIOR APPLICATION NUMBER: PCT/US01/09847			
; PRIOR FILING DATE: 2001-03-28			
; PRIOR APPLICATION NUMBER: US 60/193,446			
; PRIOR FILING DATE: 2000-03-31			
; NUMBER OF SEQ ID NOS: 1588			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1584			
; LENGTH: 36296			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 282180			
US-10-240-425-1584			
Query Match 9.0%; Score 353.6; DB 13; Length 36296;			
Best Local Similarity 57.2%; Pred. No. 8.4e-83;			
Matches 956; Conservative 0; Mismatches 609; Indels 107; Gaps 13;			
QY	522	GCTAATTTTTTTTTTTTTTAAATTTTCATTTATGTATTCATTTTCTTTCTTTTGTGTGT 581	
Db	15559	GGTAATGTTGAGCTTCTCAAGCAGCATTTATCTATTTATTATTCATGATTTATT 15500	
QY	582	TGTTTTGAGATAGGTTCTACTTTGTTTACCAGGCTGGAGGGCAGTGGCATGTCACAGC 641	
Db	15499	TTTTTTAGACAGAGGTTCTCATTCCTCCACCGGCTGGAGTAGTTGTGCAATCATAGCTC 15440	
QY	642	TGACAGCCTTGACTTCTCGGGCTCAAGTATCCTCTCGCTCAGCCTCCCAAGTAGCTG 701	
Db	15439	ACTGCAGCCTCCAACTCCTGGGCTCAAGTGATCCTCACACCTCAGGCTCCTGAGTAGTTG 15380	
QY	702	GGACTCAAAACAGCTGTCAACATGSCCTGGCTGATATTTTTTTCTTGAAACAGGGTATCA 761	

14381 AACAGAAACCTGCTGTCTCATATGCATATATGCAACAACAACAACAACAACAACAACAACAACA 14322
 QY 1833 GAAATAAGAATGAATAAATAATTCTCCG-----GGCCAGGCGTGGTGCTTCATGCCCTGTAA 1884
 Db 14321 CACTGTACATGAACTTACACAACAAGACATCAAGAGCAGCGGCGTGGTGCCTCACACCTGTAA 14262
 QY 1885 TCCAGACACTTTGGGAGGCCAAGCGATNGTGATCACTTAGGTCTAGGAGTTCAAAACCAG 1944
 Db 14261 TCCTAGCACTTTGGGAGGCCAGGTGGGTGGATTGCTCAGGTCTAGGGATTGGAGACCAG 14202
 QY 1945 CCTGGCCAACATGGTCGAACCCCCGCTCTACTATAAAAAAACAAAAAATAAAAAAATAGCTTAAGTGTGG 2004
 Db 14201 CCTGGCCAACATAGTGAACCCCCCATCTCTACTATAAAAAATAC-AAAAAATAGCTTGGCATGG 14143
 QY 2005 TGCGCATGCTCTTAATCCCAGCTACTTGGGAGGGTGAGGACAGGAGAAATTTCTTGAACCC 2064
 Db 14142 TGCGGGTGCTGTATCCAGCTACTCGGAGGCTGAGGCAAGAGAAATCGCTTGAACCC 14083
 QY 2065 GGGAGGACAGGTTGCAGTG-----AAGCGAGATCACACCACTGCCTCCAGCCTGGGG 2118
 Db 14082 AGG-GGTGGAGGTGCACCTGAGGCCAAATTGAAATTGAACCACTGCCTCCAGCCTGGGT 14024
 QY 2119 GAGAGGCGAGACTTCTCTCAAAAAACAAAAACAAAGAAATTAAGCAA 2170
 Db 14023 GAAAAAGCAAACTCCGTGTAGAAAACAAAACGAAACAAAAACCCAGA 13972

 RESULT 10
 US-10-087-192-52
 ; Sequence 52, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: CANCER
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 52
 ; LENGTH: 60815
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(60815)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-087-192-52

 Query Match 9.0%; Score 353; DB 13; Length 60815;
 Best Local Similarity 55.0%; Pred. No. 1.7e-82;
 Matches 1044; Conservative 0; Mismatches 710; Indels 145; Gaps 12;

QY	328	TATATCTCAAAATGGATTCACCTGTGTGGGTATCCCATATCATATATTATAGAGACAGTT	387
Db	4423	TTGTGTTAAACAGCAACTTTCTTTTTCTTTTTCTTTTTCTTTTTTTTAAAGATAGGGT	4482
QY	388	CTCACTCTGCAACCCAGGCTGGAGTGGCAATCATAGCTCACTGTAACTCAAA	447
Db	4483	CTCACTCTCATGCCCATGTTGGAGTGCATGTCATGCTTTGGCTCATTCAGAGCTCC--	4540
QY	448	CTCCTGGGCTCAAGTGATCCTACTACCTCAGCCTCCAGAGTAGCTACAGGCACA	507
Db	4541	-----AGGAGTCTCTCCACCTCAGCCTCTGAGAGCTTGAGCCACAGGCACA	4588
QY	508	CACAGCCATACCTGGCTAAATTTTTTTTTTAATTTTCAATTTTANGTATTCATTTTC	567

D	b	4589	GTCACCATGCCCACTTAAATTTTTTTTGAGATGGGTC-----	4623
Q	y	568	TTTTTTGTGTGTGTGTGAGATAGGGTCTCAGCTTTGTACCCAGGCTGGAGGCAGT	627
D	b	4629	-----GTCCTTGCTCTGTACACCAGGCTGGAGTGCAGT	4660
Q	y	628	GGCANGGTGACAGCTGA--CGACGCTTGACTCTCTGGGCTCAAGTGTATCTCTCGCTCA	685
D	b	4661	GGCANGATCTCCGCTCACTGCTCCCCTGCTCTCTGGGCTCAAGGATCTCTCCACCTCA	4720
Q	y	686	GCCTCCCAAGTAGTGGGACTACAACACAGTGTCAAGTGCCTGGCTGCTGATATTTTTTTC	745
D	b	4721	GCCTCCCAAGTAGTGGGCTACAGGCACAAGCATGGAGCCACAGCTAATTTTTTCTATT	4780
Q	y	746	TT-----GAAAACAGGGTATCACCTCTGTGCCCCAGGCTGGAGTACAGTGGCGTAATAATAG	800
D	b	4781	TTTGGCACAGCAGGGTTTCAACATTTTGCCCCAGGCTGGTCTCAA-----	4825
Q	y	801	CTCACTGACGCTCCCTCTCTGGGCTCAAGCAATCCGCTGGCCTCAGCATCTCTGAGTAGC	860
D	b	4826	-----ACTCTGAGCTCACTGATCTCCACCTGGGCTCAAAAGTGC	4870
Q	y	861	TGGGACTACAGGCTGTGCCCCAGGCCCAGCTAAAGTTTAAAAAAGATTTTGTGTATA	920
D	b	4871	CAGGAITGCAAGTTTGAGCCACCGCACTGGCCCTAAATTTTGTATTTTTTTTTTTTTT	4930
Q	y	921	GAGGAGTCTTGCTATGTGCTCAGGCTGTATTTTATTGTTGAGACAAGGCTCACTAT	980
D	b	4931	TTTTGTAGATGGGTTGCGCCATGTGCCCCAGGCTGGTAAACAGCAATTTTCTTCT	4990
Q	y	981	GTGGCATGATCCCCCACTCCCACTGCCAAAGTCTCATCTATCTGTTCATTAGTCA	1040
D	b	4991	TCCTCTTTTTTCTGAGATGGAGTTTAGCTCTTGGAGTGTGATGGCATGATCTGGCTCA	5050
Q	y	1041	GTTCACAGCAATTTAGTTGTTTTCACATTTTGGACCATTTAGTAATTAATCTCCAGTGAAT	1100
D	b	5051	CTCAACCTCTGCTCTCTGGTTTCAAGCAATCTCC---TGCTCATCTTCCAAGTAG	5106
Q	y	1101	AATCATGTATACATTTGTGGGCATATGTTTTCAATCTGTGTGGTATTATCTAGGAG	1160
D	b	5107	CTGGGATACAGGTGCTGCCCCCACCCAGCAATTTTTTTGTCTTTAGTAGATGG	5166
Q	y	1161	TGGAATGTCTGGATCCGGGTAAATTTTTCACAGGCAGAGTTCAGGGGA-----AGAA	1213
D	b	5167	GGCTTCATCATATTTGGCCAGATGGTCTTGAATCTCTGACCTCAGGTGATCCGCCATCT	5226
Q	y	1214	AAACTTGGGAAATGAAGCATGTTTGAATATGACGAGAGTGCAGGGGTTTTTCGAGTT	1273
D	b	5227	CAGCCTCCCAAGTGTGGGATTTATAGTGTGAGCCACCATGCTGGCCTTTTCTTCTT	5286
Q	y	1274	TTATTTTATATCTGTGTGACAAATGT---GCAGTTTGTGAAGATACAAGTTATATAAG	1330
D	b	5287	TCTTTTTTTTTTTTTATAGCAGGGTCTGCTTGTFCACCCAGACTGGAGTGCAGTGCCA	5346
Q	y	1331	TGAAAGTGAAGATTAAGGCTTGAATPAGGCGGTTTCAAGATAAAATCATGAAGCATTTGA	1390
D	b	5347	CCATCTCGGCTCATTAACAACCTCCACTCCAGTATCAAGTAAATCTCGTGCCTCAGCCT	5406
Q	y	1391	ATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATAAAAAATCACAATTTTTTTTT	1450
D	b	5407	CTCCAGTAGCTGGGACTACAGGTGCCACCGCAGGCTCTGGTCTGGCTTTTTTTTTTT	5466
Q	y	1451	TTTTTTTGAGAAAGAGTCTTGCTCTTTTACCTCGGCTGGAGGCGAGTGTGATCTCAGC	1510
D	b	5467	TTTTTTTGACACAGAGTCTCGCTGTGCCCCAGGCTGGAGTGCAGTGTGTGATGATAGC	5526
Q	y	1511	TCACTGCAATTTTCGCTCCCGGTTTCAAGCAATTTCTCTGCTTCAGCCTCCCAAGTAGC	1570
D	b	5527	TCACTGAGCTCCAACTCCCTGACTCAAGCAATCTCCAGCCTCAGCCTCCCGAATAGC	5586
Q	y	1571	TGGGACTACAGGCATTTCCCAACAATGCCAGCTGTATTTTTTGTATTTTATAGTAGATGG	1630
D	b	5587	TGGGATTAAGGCATGAGGCAACATGCCAGCTAATTTCTGTATTTTATAGTACACAGG	5646

Qy	1631	ATTTCAC	TTTGTG	TGGCC	CAAGTGG	GTCTCAAC	-----	TTTT	1666
Db	5647	GTITCT	CCATGT	TGGC	CAGGCT	GTGTCT	CGAACT	CTTG	5706
Qy	1667	TGCTGT	CATAA	TTTGTG	TAACTA	TATG	TTCTCT	TTTG--	1724
Db	5707	AGCCT	CCCAA	GTCT	GGGATTA	CACCG	CGCTTG	CCCTTAA	5766
Qy	1725	AAAAA	TAAAT	CTTTAG	AAAT	COAA	TACAGT	GTGTGG	1784
Db	5767	TGGCG	GTGTAG	CAGC	AGCAT	CAGAG	GTGTCT	TTCTG	5826
Qy	1785	ACAGT	GTGAC	CAGG	----	GCCT	CAGG	AGTAG	1839
Db	5827	CGAGT	CCAT	CCAG	ATAG	TAGCT	CA	CAGG	5886
Qy	1840	AATGA	AAAT	ATCT	CCGG	CC	CAGCG	GTGTGG	1899
Db	5887	AA	TATAT	AG	CAAT	GAAG	CCG	GTGTGG	5946
Qy	1900	AGGCA	AGGC	ATGTGG	AT	CACT	CAGT	GTGAG	1959
Db	5947	AGGCT	GAGG	CGGTGG	GTG	CACT	TAAG	GTGAG	6006
Qy	1960	GAAC	CCCC	GTCT	CTTACT	AAAA	ATA	CAAAAA	2019
Db	6007	GAAC	CCCC	GTGTCT	ACT	AAAA	ATAC	AAAA	6065
Qy	2020	ATCC	AGCT	ACTT	GGG	GGGTG	AGG	CGTGTG	2079
Db	6066	ATCC	AGCT	ACTT	CGAG	GGGTG	AGG	CGTGTG	6125
Qy	2080	CAGT	GAAG	CGAG	ATCA	CA	CCACT	GCAT	2139
Db	6126	CGGT	GCCT	GAG	ATTAT	TG	CCACT	GCAT	6185
Qy	2140	AAAAA	CA	AAAAA	CA	AAAG	AA	TTAG	2178
Db	6186	AAAAA	CA	AAAAA	CA	AAAG	AA	TTAG	6224

RESULT 11

US-10-181-875-10

; Sequence 10, Application US/10181875

; Publication No. US2003021633A1

; GENERAL INFORMATION:

; APPLICANT: Isis Pharmaceuticals, Inc.

; APPLICANT: Brett P. Monia

; APPLICANT: Robert McKay

; APPLICANT: Madeline M. Butler

; APPLICANT: Jacqueline Wyatt

; TITLE OF INVENTION: ANTISENSE MODULATION OF GLYCOGEN SYNTHASE KINASE 3 ALPHA

; FILE REFERENCE: RTSP-0356

; CURRENT APPLICATION NUMBER: US/10/181,875

; PRIOR FILING DATE: 2002-07-22

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 88

; SEQ ID NO 10

; LENGTH: 12394

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (115)...(397)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (2438)...(2625)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (5639)...(5722)

; FEATURE:

[illegible]

Db 4214 GAGGACGAGAAATCTTGAACCCAGGAGCGAGGTTCCAGTGCAGCCAGATCACACCA 4273

Qy 2101 CTGCACTCCAGCTCGGGGAGAGAGAGACTTCCTCTCAAAAAACAAAAACAAAAGA 2160

Db 4274 CTGCACTCCAGCTCGGGGAGAGAGAGTGGCTGTGCTCTCAAAAAACAAAACAAA 4333

Qy 2161 ATTAAGCAAAATTAGACA 2177

Db 4334 ACAAAACAAAACAAAACA 4350

RESULT 13

US-09-764-891-7809

; Sequence 7809, Application US/09764891

; Publication No. US20030077808A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC006

; CURRENT APPLICATION NUMBER: US/09/764,891

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 10231

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7809

; LENGTH: 29163

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-891-7809

Query Match 8.9%; Score 347.4; DB 10; Length 29163;

Best Local Similarity 69.9%; Pred. No. 3.4e-81; Indels 11; Gaps 3;

Matches 515; Conservative 0; Mismatches 211;

Qy 1441 TTTTCTTTTCTTTTGTGAGAAAGAGTCTTGCTCTTTCAACCTGGCTGGCTGGAGGGCAGTGGTG 1500

Db 3623 TCTTTCTTTTCTTTTCTTGAGTCTTGCTTTGTCGCCAGGCTGGAGTGAGTGGTG 3682

Qy 1501 TGATCTAGCTACTGCAATTTTGGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGCTT 1560

Db 3683 CGATCTCAGCTCACTGCAAACTCCACTCCCGGGTTCAAGAGATTTCTCTGCTTAGCT 3742

Qy 1561 CCCAGTAGCTGGAGTACAGGCACTTCCACCATGCCAGCTGATTTTGTATTTTATAG 1620

Db 3743 CCTGAGTAGCTGGAGTACAGGCTGCACTACCGTCCCTGGCTAAATTTTATATTTTAA 3802

Qy 1621 TAGAGATGGAGTTTCACTTTTGTGGCAAGCTGGTCTCAAACTTTTGTGTCATATTTG 1680

Db 3803 TACAGATGGGTTTCACTTTGTTGGTCAAGCTGGTCTCGAACTCTGACTTCATGATCCG 3862

Qy 1681 TTGTAACTATTTCTTTTCTGAGTAGGGGGCCCCAGACCAAAAAAATAATCTTAG 1740

Db 3863 TCCCCCTCGACCTCCCAAAAGTGTAGGATTAAGGATGAGCCACTGTGCCAGCC---- 3918

Qy 1741 AATCCAAATCAGTGTGTGGTTTGACACTGTCTACATTGAGAACCAAGTGTGACCAAGGC 1800

Db 3919 ---GCTAAATTTTGTACAGACAGAGTTTCTCCATGTTGCCAGGCTGATCTGGAA 3975

Qy 1801 CTCAGGAGTAGAGTGTATCTCTCTCGAAGAGAAATAGAAATAATTTCCGGGCC 1860

Db 3976 CTCCTGGGCGAAGTGAACCATCTGCTCGC---CTCCAAAGTGTGGGATTAAGGCC 4032

Qy 1861 AGCGGTGGTGGCTATGCTGTAATCCAGCACTTTGGAGGCCAAGGATGTGATCTAC 1920

Db 4033 GGGTGGGTGGCTCAGCCCTGTAATTCAGCACTTTGGAGGCCAAGGAGTGGATCAG 4092

Qy 1921 CTGAGGTGAGGAGTTCAAAAACAGCTGGCCACATGTTGAAACCCCGTCTCTACTAAA 1980

Db 4093 CTGAGGTCCGAGGTTTAAGACAGCCCTGGCCAAAGAGTGAACCCCTCTCTACTAAA 4152

Qy 1981 ATACAAAAATTAGCTAAGTGTGGTGGCGCATGCTGTATCTCCAGCTACTTGGAGGGT 2040

Db 4153 ATAC-AAAAATAAGCCAGGTGTGGTGGCGCATGCTGTATCTCCAGCTACTTGGAGGGT 4211

Qy 2041 GAGGACGAGAAATTTCTTTGAACCCAGGAGCGAGGTTGCAGTGAAGCGAGATCACACCA 2100

Db 4212 GAGGACGAGAGAAATCACTTTGAACCCAGGAGCGAGGTTGCAGTGAAGCGAGATCACACCA 4271

Qy 2101 CTGCACTCCAGCTCGGGGAGAGAGAGACTTCCTCTCAAAAAACAAAAACAAAAGA 2160

Db 4272 CTGCACTCCAGCTCGGGGAGAGAGAGTGGCTGTGCTCTCAAAAAACAAAACAAA 4331

Qy 2161 ATTAAGCAAAATTAGACA 2177

Db 4332 ACAAAACAAAACAAAACA 4348

RESULT 14

US-09-880-107-2278

; Sequence 2278, Application US/09880107

; Patent No. US20020142981A1

; GENERAL INFORMATION:

; APPLICANT: Horne, Darci T.

; APPLICANT: Vockley, Joseph G.

; APPLICANT: Scherf, Uwe

; TITLE OF INVENTION: Gene Logic, Inc.

; FILE REFERENCE: 44921-5028-WO

; CURRENT APPLICATION NUMBER: US/09/880,107

; CURRENT FILING DATE: 2001-06-14

; PRIOR APPLICATION NUMBER: US 60/211,379

; PRIOR FILING DATE: 2000-06-14

; PRIOR APPLICATION NUMBER: US 60/237,054

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 3950

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2278

; LENGTH: 26928

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M17262

US-09-880-107-2278

Query Match 8.8%; Score 342.8; DB 9; Length 26928;

Best Local Similarity 57.5%; Pred. No. 5.5e-80;

Matches 848; Conservative 0; Mismatches 552; Indels 74; Gaps 10;

Qy 735 TATTTTCTTTTCTTGAAAACAGGGTATCATCTCTTCCAGAGCTGGAGTACAGTGGCGTAA 794

Db 3122 TTTTCTTTTCTTGGAAGTCAAGAGTCTTGCTCTGTGCCAGGCTGGAGTCAAGTACCA 3181

Qy 795 TAATAGCTCACTGCAGCCCTCC--CCTCCTGGGCTCAAGCAATCCGCTGGCCTCAGCATCT 853

Db 3182 TCTAGCTCACTGCAGCCCTCCGCTCCCGGTTCAAGCTATTCTCTCTGCTCAGCCTCC 3241

Qy 854 GAGTAGCTGGGACTCAGGCTTGTGCCACAGCCAGCTAAGTTTAAATAATGATTTT 913

Db 3242 AAATAGCTGAGACTATAGGCACGCATCTCATGCCAGCTAATTTT-----TTATTTT 3295

Qy 914 TGGTATAGAGAGGCTTGTGCTATGTTGCTCAGGCTGATTTTATTGTTGACACAAGGTC 973

Db 3296 TAGTAGAGACGAGGTGTCTCCATGTTGGCCAGGTTGGTCTTGAAC----- 3341

Qy 974 TCATATGTTGCCATGATCCCGCCACCTCCACTTCCAAAGTGCATCTTATCTGTTC 1033

Db 3342 --CCTGTCTCAAGTATCCACTGCTCGGCTCCCAAGTGGGATTCAGGCAATG 3399

Qy 1034 TTAGTCAAGTACAGACATTTAGTGTGTTTCCACTTTTGGACCATATGAAATAACTCC 1093

Db 3400 AGACACCGCGCCCGGCC-----TGCTTGTCCCTTCTTAAATAGTGTGCTCCATTTGTA 3454

Qy 1094 AGTGAATATTATGATATATATGTTGTTGGCATATGTTTTCATTTCTGTTGGTTTATAT 1153

Db 3455 GCTGCTGATTTCTTTGGGACATTTGCTCCGTAACCTTTTCATAAGCATCAGTGATTCA 3514

1154	QY	CTAGGAGTGGAAATGCTGGATCCCGGGTAATAATTTTGACAGGACAGNGTTCAGGGGAGAA	1213
3515	Db	CCATTCTTCCACCCAAAGCTTCACCGTAAATTTGTTGTTCTTGCTTCAATTTTCAGCA	3574
1214	QY	AAACTTGGGAAAATGAAGCATGTTTAGAATCAGCAAGAGTGCAGGGGTTTTTCGGAGTT	1273
3575	Db	GAATTCATTAGCTCTGTGATAAGGGCTCGCTTCAAACCTGATGCTTATCCTTCTTAGTGCC	3634
1274	QY	TTATTTTATATTTCTGTGTGACAAATGTC-----AGTTTGATGAAGATACAAGTTTACTTA	1328
3635	Db	TCAAACTACATCTGTCTCATGTTATATAGCAAGTTAGTGTGAGTTTATTTTGTGTGCAC	3694
1329	QY	AGTGAAGTGAAGAAATTAAGGCTGGATAGGGCGTTTCAGAGTAAATCATGAGAGCACTTT	1388
3695	Db	AAAAATTTTTTTTAAATCCCATGCACTTTTTTTCATTAATACGCAATTTTCCATGACATTTTC	3754
1389	QY	GAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAAATAATAAAAAATCAATTTTTTTTT	1448
3755	Db	GAAGACCCCTTGTAGATGCTGTGTTGTTTAAACCAACCAGTTTA---CAGTAATTTTTTT	3810
1449	QY	TTTTTTTTTGAGAAAGAGTCTTGCTCTTTCACCTCGCTGGAGGGCAGTGTGTGATCTCA	1508
3811	Db	TTTTTTTTTGAGATCAAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCATTTGGCACACTCTCG	3870
1509	QY	GCTCACTGCAACATTTGCGCTCCCGGGTTCAAGCAATTTCTCCTCTTCAGCTCCCAAGTA	1568
3871	Db	GCTCACATGCAACCTTCGCTCTCTGGGTTCAAGCAATTTTCTCTCAGTCTCCCGAGTA	3930
1569	QY	GCTGGGACTACAGGCACTTCCCACTGCCCAGCTGATTTTTTGATTTTTTAGTAGAGATG	1628
3931	Db	GCTGGGATTCAGGCTGTGTGCCCACTGCTAGCTAATTTATGTGTTTTTAGTAGAGACG	3990
1629	QY	GGATTTCACTTTCTGGCCAAGCTGTCTCAAACTTTTTCCTCTTCATAATTTGTTGTAAC	1688
3991	Db	GGGTTTCACTATGTTGGCTAGGCTGGTCTCGAACTCCTCACTTGTGATCGGCCCGCTC	4050
1689	QY	ATTGTTCTTTTTCTGAGTGTAGGGCCCCCAGACCAAAAAATAATCTTTAGAAATCCAAA	1748
4051	Db	GGCTTCCCAAAGTATTGGGATTACAGCGTGA-----GACTCTTGCACTTTGGCC	4099
1749	QY	TCAGTGTGTGTTTGTACCACTGTCTCTAGAAACCAAGTGTGACCGGCTCTCAGGAG	1808
4100	Db	TACAGTAATTTTATAGCAGCTTAGGCTTAAGATAGCCATTTCTGGGTATAGAATGTCA	4159
1809	QY	TAGAGGTGATCTCTGTCTCGAAAGAGAAATAGAAATGAAATATTCTCGGCCACAGGGTGG	1868
4160	Db	TACTGAACAGCCCTGCAACTGTGATGATAAAGTCTGCAAGAGGCC-----GGGCAG	4210
1869	QY	TGGCTCATGCTGTATPCCCAGCACTTTTGGAGGCCCAAGCAGTGTGATCACTGAGGTC	1928
4211	Db	TGGCTCATATCTGTAACTCCACAGCACTTTGGGGGGCCGAGGACAGTGGATCACCTGAGTC	4270
1929	QY	AGAGTTCAAAAACAGCTCGCCCAACATGCTGAAACCCCGCTCTCTACTTAAAAATACAAAA	1988
4271	Db	AGCAGTTTGAGACGAGCTGACCAATGTGTGAAACCCCACTCTCTACTAAAAATAC--AA	4328
1989	QY	AATTAGCTAAGTGTGTGGCGCATGTCCTGTAATCCCC-----AGCTACTTG	2033
4329	Db	AATTAGCTGGGCGTGTGTGATGTCATGCTTGTAAATCCCTAGCATGCACCTTGGGAGCTACTTG	4388
2034	QY	GGAGGTTGAGCAGGAGAAATTTCTTGAAACCGGGAGCAGAGTTGCAGTGAAGCCGAGAT	2093
4389	Db	GGAGGCTGAGCAGGAGAAATCACTTTGTACTCAGAGGCCGAGGTTTGCAGTGTGAGT	4448
2094	QY	CACACCACTGCATCTCCAGCTTGGGGGAGAGAGCGAGACTTCTCTCAAAAAAACAACAAA	2153
4449	Db	CACGCCACTGCATCTCTTTCTGGGTGACAGAGTGTGAGACTCCATCTCAAAAAACAACAA	4508
2154	QY	CAAAAGAAATTAGCAAAATTAGCATTCGACAGAG	2187
4509	Db	AAAAAACAACAAAAACAACAAAAAACCACACAG	4542

RESULT 15

US-10-235-192A-34
; Sequence 34, Application US/10235192A
; Publication No. US20040043389A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; TITLE OF INVENTION: Methods and Compositions for Identifying
; TITLE OF INVENTION: Risk Factors for Abnormal Lipid Levels and the Diseases
; TITLE OF INVENTION: and Disorders Associated Therewith
; FILE REFERENCE: MMI-011
; CURRENT APPLICATION NUMBER: US/10/235,192A
; CURRENT FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 26928
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-235-192A-34

Query Match 8.8%; Score 342.8; DB 13; Length 26928;
Best Local Similarity 57.5%; Pred. No. 5.5e-80;
Matches 848; Conservative 0; Mismatches 552; Indels 74; Gaps 10;

735 TATTTTTTTTCTTTGAAACAGCGGTATCACTCTGTTTGCCCGAGCTGGAGTACAGTGGCGTAA 794
3122 TTTTTTTTTTGGGAGTCAAGAGCTTGTCTCTGTTTGCCAGGCTGGAGTGCAGTGACACGA 3181

795 TAATAGCTCACTGAGCGCTCC-CCTCTGGGCTCAGCAATCCGCTGGCTCAGCATCCT 853
3182 TCTCAGCTCACTGAGCGCTCCGGCTCCGGGTTCAGCATATTCTCCGTCTCAGCCTCCC 3241

854 GAGTAGCTGGGACTACAGGCTTGTGCCACAGGCCCAAGCTTAAGTTTTTAAAAATGATTTT 913
3242 AAATAGCTGAGACTATAGGCAGGCACATCATGCCAGCTAAATTTT-TTATTTT 3295

914 TGGTATAGAGAGGTCTGTGATGTTGCTCAGGCTGTAATTTTATTTGTTGAGACAAGTTC 973
3296 TAGTAGAGACGAGGTGTCTCATGTTGGCCAGGTGTGCTTTGAAC-T----- 3341

974 TCATATGTTGCCATGATCCCCACCTCCACTTCCAAAGTGCTCATCTTATCTGTTCA 1033
3342 --CCTGTCCTCAAGTGATCACTGCTCCGCTCCCAAGTGTGGGATTCAGGCATG 3399

1034 TTAGTCAAGTTGACAGACATTTAGGTTGTTTCCATTTTTCAGCATATGATTAATCTCC 1093
3400 AGACACGCGCCCGGCC-----TGCTGTGTCCTCTTAAATAGTGTGCCATTTGTAA 3454

1094 AGTGAATATTCATGTATACATTTGTGGGCATATGTTTTCATTTCTGTGGGTTTATAT 1153
3455 GCTGCTGATTTCTTTGGGACATGTCTCCGTAACATTTTCATAAGCATCAGTGATTTCA 3514

1154 CTAGGAGTGGAAATGCTGGATCCCGGGTAATATTTTCAGGCGAGGTTTCAGGGGAAGAA 1213
3515 CCAATCTTCCACCCCAAGCTTCAACGTAATTTGTTGTTGTTCTTCTGCTCAATTTTCAGCA 3574

1214 AAACCTGGGAAATCAGCATGTTTGAATATCAGCAGAGTGCAGGGGGTTTTTCGGAGTT 1273
3575 GAATTCATTTAGCTCTGATAGGGCTCGCTTCAAACTGATGTCTTATCTCTTTAGTGCC 3634

1274 TTAATTTATATTCGTTTGAACAAATGTGC-----AGTTTGATGAAGATACAAGTTTATACTA 1328
3635 TCAAACTACATCCTGTTTCACTCATGTTATAGCAAGTTAGTGTAGTTTATTTTGGTGCAC 3694

1329 AGTGAGAAGTGAGAAATTAAGCTCGGAATAGGGCGTTCAGAGTAAATCATCAGACATTTT 1388
3695 AAAAATTTTTTAAATCCATGACGTCTTTTTTTCATAATACGCATTTTCCATGAACTTTTC 3754

1389 GAATACCAAAATTAAGGAGCTTGCTGTAAACAAATAATAAAAAATCACAAAATTTTTTTT 1448
3755 GAAGACCCCTGTAGATGTCTGTGTTTAAACCCAGTTTA-----CAGTAAATTTTTTTT 3810

1449 TTTTTTTTGAAGAAGCTTGTGCTCTTTTACCCTGGCTGGAGGCGAGTGGTGATCTCA 1508

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:19 ; Search time 14370.8 Seconds

(without alignments)
8124.892 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3910

Perfect score: 3910

Sequence: 1 aggaatcattccatttaaat.....cggacacggtcaccgctgac 3910

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*

5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*

9: gb_estl.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*

13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*

17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*

21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*

25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	620.4	15.9	1238	12	BM563325 AGENCOURT
2	491.4	12.6	579	12	BI559696 603252566
3	450.2	11.5	471	28	AQ760013 HS_3202_A
4	294.2	7.5	1910	11	BC035771 Homo sapi

5	290.6	7.4	2076	11	BC035510
6	290.6	7.4	2772	11	BC038630
7	286	7.3	754	12	BI827092
8	283.4	7.2	2772	11	BC038630
9	273.4	7.0	2076	11	BC035510
10	270	6.9	749	13	BU175345
11	269.4	6.9	1910	11	BC035771
12	259.4	6.6	929	9	AL580585
13	259.2	6.6	693	13	BU616025
14	255.8	6.5	2971	28	AF101960
15	255.2	6.5	958	9	AL578531
16	254.4	6.5	764	13	BU617735
17	252.4	6.5	940	13	BX424350
18	249.4	6.4	524	10	AW973992
19	249.2	6.4	580	12	BQ053726
20	249	6.4	978	13	BX325356
21	248.4	6.4	1201	13	BX385602
22	248.2	6.3	731	9	AI687343
23	247.2	6.3	775	12	BM905333
24	246.8	6.3	521	13	BU617227
25	246.8	6.3	598	12	BQ017808
26	246.8	6.3	616	13	BU617236
27	246.2	6.3	508	28	AQ418545
28	246.2	6.3	958	9	AL578531
29	246.2	6.3	969	13	BX415714
30	245.6	6.3	2284	11	BC036238
31	245.4	6.3	467	28	AQ388230
32	245.4	6.3	785	12	BM721013
33	244.8	6.3	663	29	AG174449
34	244.6	6.3	750	12	BM678642
35	244.2	6.2	1201	13	EX385602
36	244	6.2	548	14	CB144391
37	243.8	6.2	755	28	BZ599689
38	243.8	6.2	2971	28	AF101960
39	243	6.2	826	28	BZ611620
40	243	6.2	929	9	AL580585
41	242.6	6.2	493	10	AW328331
42	242.4	6.2	532	12	BG830539
43	242.4	6.2	861	12	BG676827
44	242.2	6.2	390	13	BU588888
45	242.2	6.2	759	13	BQ436286

ALIGNMENTS

RESULT 1
BM563325
LOCUS
DEFINITION AGENCOURT_6564864 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5743936
5', mRNA sequence.
ACCESSION BM563325
VERSION BM563325.1 GI:18810153
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1238)
NIH-MGC <http://mgs.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbe-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL12764 row: h column: 17
High quality sequence stop: 622.

RESULT 2
B1559696
LOCUS B1559696 579 bp mRNA linear EST 05-SEP-2001
DEFINITION G03232565F1 NIH MGC 97 Homo sapiens cDNA clone IMAGE:5294921 5'

```

Db      364 CAGCAGATTGGCGGAGCAATGGCAAGCGCAGGAGCGACCAAAATCCCTTCG 423
QY      3644 GCCAATGACGAGCGGAGTTTACAGAGCCTCATTAGCATTTCCCGAGAGC-AGGGGCA 3702
Db      424 GCCAATGACGAGCGGAGTTTACAGAGCCTCATTAGCATTTCCCGAGAGCTAGGGGCA 483
QY      3703 GGGCAGAGGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3762
Db      484 CGGCAGAGGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 543
QY      3763 TCGCGGAGAGCTCGGCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3798
Db      544 TCGCGGAGAGCTCGG-CTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 578

RESULT 3
AQ760013
LOCUS
DEFINITION
  HS_3202_A1_B01_MR_CIT_Approved_Human_Genomic_Sperm_Library_D_Homo
  sapiens_genomic_clone_Plate=3202_Col=1_Row=C, genomic survey
  sequence.
ACCESSION
  AQ760013.1 GI:5625489
VERSION
  93980589
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 471)
  Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
  Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
  Hood,L.
  Sequence-tagged connectors: A sequence approach to mapping and
  scanning the human genome
  Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  10449764
COMMENT
  Contact: Mahairas GG, Wallace JC, Hood L
  High Throughput Sequencing Center
  University of Washington
  401 Queen Anne Avenue North, Seattle, WA 98109, USA
  Tel: (206) 616-3618
  Fax: (206) 616-3887
  Email: jwallace@u.washington.edu
  Clones may be purchased from Research Genetics (info@resgen.com).
  BAC end Web Server: http://www.htsc.washington.edu
  Plate: 3202 row: C column: 1
  Seq primer: M13 Reverse
  Class: BAC ends
  High quality sequence stop: 471.
  Location/Qualifiers
    1..471
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone="Plate=3202_Col=1_Row=C"
      /sex="male"
      /clone_lib="CIT Approved Human Genomic Sperm Library D"
      /note="Organ: sperm; Vector: pBelobAC11; BAC clones in
        E-Coli DH10B"

FEATURES
  source
    1..471
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone="Plate=3202_Col=1_Row=C"
      /sex="male"
      /clone_lib="CIT Approved Human Genomic Sperm Library D"
      /note="Organ: sperm; Vector: pBelobAC11; BAC clones in
        E-Coli DH10B"

ORIGIN
  Query Match 11.5%; Score 450.2; DB 28; Length 471;
  Best Local Similarity 97.0%; Pred. No. 2.1e-59;
  Matches 455; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      9 ATCCATTAAATCATACAATTAATGGCTTTTAGTATATTACAGGTTGTGCATCCATCA 68
Db      1 ATCCATTAAATCATACAATTAATGGCTTTTAGTATATTACAGGTTGTGCATCCATCA 60
QY      69 CAATCCATTTAGAACAGTTTATTACTCCAAATAAACCTGCATCTCTAGCATCA 128

```

```

Db      61 CAATCCATTTAGAACAGTTTATTACTCCAAATAATAACCTGCATCTCTAGCATCA 120
QY      129 CCCCCCAACATCTCTCCATCTCTCCATCTCCATCTCCATCTCCATCTCCATCTCTCTCTCT 188
Db      121 CCCCCCAACATCTCTCCATCTCTCCATCTCCATCTCCATCTCCATCTCTCTCTCTCT 180
QY      189 TCTATAAATTTGCCAATCTTGGACATTTTCATATATAATGGAAGCAACAACATGTGAGACT 248
Db      181 TCTATAAATTTGCCAATCTTGGACATTTTCATATATAATGGAAGCAACAACATGTGAGACT 240
QY      249 TTGTGACTGGCTGCTTTTACCTAGCATTTCTATTTTAAAGGCTCATTATGTTACAGTACTT 308
Db      241 TTGTGACTGGCTGCTTTTACCTAGCATTTCTATTTTAAAGGCTCATTATGTTACAGTACTT 300
QY      309 AGCAGTACTTCAATCTTTTATTTCTCAATGGTATTTCCACTGTGTGGGTATCCCAATATC 368
Db      301 AGCAGTACTTCAATCTTTTATTTCTCAATGGTATTTCCACTGTGTGGGTATCCCAATATC 360
QY      369 ATATTATTAGAGACAGTTCTCTACTCTGTCCAGAGCTGGAGTGCAGTGCACAATCAT 428
Db      361 ATATTATTAGAGACAGTTCTCTACTCTGTCCAGAGCTGGAGTGCAGTGCACAATCAT 420
QY      429 AGCTCACTGTAACTCAACTCCTGGCTCAGTGATCCTACTACTCTCA 477
Db      421 AGCTCAACTGTACCTCANACTCCTGGGCTCAAGTGATCCTACTACTCTGA 469

RESULT 4
BC035771/c
LOCUS
DEFINITION
  Homo sapiens RAD1 homolog (S. pombe), mRNA (cdna clone
  IMAGE:5582524), with apparent retained intron.
ACCESSION
  BC035771
VERSION
  BC035771.1 GI:23958899
KEYWORDS
  HTC.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 1910)
  Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
  Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
  Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
  Hopkins,R.F., Jordan,H., Moore,T., Max,A.S.I., Wang,J., Hsieh,F.,
  Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
  Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
  Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
  Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
  Abramson,R.D., Mullihy,S.J., Bosak,S.A., McEwan,P.J.,
  McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
  Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
  Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
  Fahey,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Shevchenko,Y.,
  Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
  Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
  Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
  Butterfield,Y.S., Krzywinski,M.I., Skalski,U., Smalios,D.E.,
  Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  Generation and initial analysis of more than 15,000 full-length
  human and mouse cDNA sequences
  Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  22388257
  12477932
REFERENCE
  2 (bases 1 to 1910)
  Strausberg,R.
  Direct Submission
  Submitted (31-JUL-2002) National Institutes of Health, Mammalian
  Gene Collection (MGC), Cancer Genomics Office, National Cancer
  Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
  USA
REMARK
  NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT
  Contact: MGC help desk
  Email: cgapbs-r@mail.nih.gov

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (L1NL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland,
Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nhgri.nih.gov

Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgoun,C., Vogt,J.B., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRL Plate: 39 Row: P Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein

This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
1..2076
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5094058"
/tissue_type="Lymph, Burkitt lymphoma"
/clone_lib="NIH_MGC_8"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN

Query Match 7.4%; Score 290.6; DB 11; Length 2076;
Best Local Similarity 67.2%; Pred. No. 2.9e-35;
Matches 498; Conservative 0; Mismatches 214; Indels 29; Gaps 5;
QY 1421 AAAATAATAAATAACACAAATTTTTTTTTTTTGGAGAGAGCTTGCTCTTTTACC 1480
DB 741 ACAACGTAAGATAGATCTTTTTTTTTTTTGGAGATGGATTTTCACTCTGTAC 800
QY 1481 CTGGCTGAGGCGAGTGTGATCTCAGTCACGCAACTTTGCCCTCCCGGGTTCAAG 1540
DB 801 CAGGCTGAGTGCAATGGCGTGATCTTGCTCACCAGAACCTCCGCTCTCTGGTTCAA 860
QY 1541 CAATTTCTCTGCTTCAGCTCCAGCTGAGTGGGACTACAGCACTTCCACCATGCCA 1600
DB 861 CAATTTCTCTGCTTCAGCTCCAGCTGAGTGGGATTAAGGATGACGACATGCAACCAAC 920
QY 1601 GCTGATTTTGTATTTTGTAGAGATGGGATTTTCACTTTTGGCCAAAGTGTCTCAA 1660
DB 921 GAT-AAATTTTGTATTTTGTAGAGATGGGTTTCAACCATGTTGTCAGGCTGTCTGA 979
QY 1661 ACTTTTGTCTCATATTTGTGTAATTTGTTTCTTTTCTGCTGAGGTAGGCCCCCAGA 1720
DB 980 ACT-----CCGACCTCTGCTGATCCGATGCTCGGCTCCCAAGTGTGGGATT 1031
QY 1721 CCAAAAAAATAATCTTAGAATCAAAATCAGTGTGTTGTTGACCACTGTCACTTTGAG 1780
DB 1032 ACATGCGTGACCCCGTCCCGACATAGATGCTTTTAAATTTCTGGAGGAAAAAGCA 1091
QY 1781 AACCAAGTGTGACAGGCGCTCAGGAGTAGAGTGATCTCTGTCGAAAGAGAAATAGA 1840
DB 1092 AAGCAAAAGAACAGCTGGATATTTTAAAG-----ACTAAAAAGGAAAAACAA 1137
QY 1841 ATGAAATATTTCTCGGCGCAGGCTGGTGGCTC-ATGCCGTGTAATCCAGCACTTTGGG 1899
DB 1138 AAAGGAGATAGACAGGCCAGAGCGTGGTGCTCAACGCTGTGTAATCCAGCACTTTGGG 1197
QY 1900 AGGCCAAGGCATGTGGATCACCTGAGGTACAGAGTTCAAAAACAGCTGGGCCAACATGGT 1959
DB 1198 AGGCCGAGGAGGTGGATCACCTGAGGTACAGAGTTCAAGACAGCTGACCAACATGGT 1257
QY 1960 GAACCCCGCTCTTACTTAAATAACAAAAATTAAGCTAAGTGTGGTGGCGCATGCCGTGA 2019
DB 1258 GAAACCCCTGTTTC-----AAAATACAAAAAATTAGCTGGGCGTGGTGGCGGACCTGTG 1312

QY 2020 ATCCAGCTACTTGGAGGCTGAGCAGAGAGATTTCTTGAACCCGGAGGAGAGGTTG 2079
DB 1313 ATCCAGCTACTTGGAGGCTGAGCAGAGAGATCCCTTGAACCCAGGAGGTGGAGTTG 1372
QY 2080 CAGTGAAGCAGATCACACCTGCTCAGCTCCAGCCTGGGGAGAGAGAGACTTCTCTC 2139
DB 1373 CAGTGAAGCAGATCACACCTGCTCAGCTCCAGCCTGGGGAGAGAGAGACTTCTCTC 1432
QY 2140 AAAAAACAAAAACAAAAAGA 2160
DB 1433 CAAAAAATAAAAAAGAGATA 1453

RESULT 6

BC038630/c

LOCUS BC038630 2772 bp mRNA linear HTC 04-MAR-2003
DEFINITION Homo sapiens, Similar to hypothetical protein FLJ20489, clone
IMAGE:5263792, mRNA.

ACCESSION BC038630

VERSION BC038630.1 GI:24116283

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2772)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (15-OCT-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbe-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: amadan@systemsbiology.org

Anup Madan, Jessica Fahey, Brin Helton, Mark Ketterman, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 73 Row: 1 Column: 4

This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
1..2772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5263792"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH_MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Query Match 7.4%; Score 290.6; DB 11; Length 2772;
Best Local Similarity 69.5%; Pred. No. 2.4e-35;
Matches 508; Conservative 0; Mismatches 189; Indels 34; Gaps 7;
QY 1441 TTTTTTTTTTTTTTTGGAGAGAGCTTGCTTT-TCACCTGGCTGGGAGGAGTGGT 1499
DB 2753 TTTTTTTTTTTTTTTGGTAAGACGGAGTTTGTCTTGTCTCCAGGCTGGAGTGGATGGT 2694
QY 1500 GTGATCTCAGCTCAGTCAACTTTTCGGCTTCGGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGCC 1559
DB 2693 GCCATCTCGGCTCGTGCAAACTGCACCTCCGGGTTCAACAGATTTCTCTGCTTCAGCC 2634

```

QY 1560 TCCCAAGTAGCTGGAGTACAGGCACTTCCACCATGCCAGCTGATTTTGTATTTT 1619
Db 2633 TCCCAAGTAGCGGAGCACAGGTGGCGGCACACGCCAGCAATTTCTGTATTTA 2574
QY 1620 GTAGAGATGGGATTTCACTTTTGTGGCAAGCTGGTCTCAAACTTTTGTGTCAATA 1679
Db 2573 GTATATTGGGGTTTACCATGTTGGCCAGGCTGGTCTCAAACTCTCTGACCTCAGGTGAT 2514
QY 1680 GTTGTAACTATTGTTCTTTTGTCTGAGTAGGGCCCCCAGACCAAAATAAATTTT 1739
Db 2513 CCATCTGCCCTCAGCTCCCAAGGTGGT-----GGATTACAGGCATGAGCCACA 2465
QY 1740 GAATCCAAATCAGTGTGTTGGTTGACCACTGTCACTTGAGAACCCAGAGTGCACAGGG 1799
Db 2464 GCTCCAGCTAGATTTCTGATATATCACTATCTGCTTGTAGAGTGAAGTCTGGC 2406
QY 1800 CTTAGAGTAGAGTGTATCTCTGCTCGAAGAGAAATAGAAATATTCTCCGGGC 1859
Db 2405 TAAATTAATAGAGATGAATTTACTGAAAAGTGTATA-----TGC 2364
QY 1860 CAGCGTGGTGGCTCATGCTGTAAATCCAGCACTTTGGAGGCCAAGGCATGTGATCA 1919
Db 2363 CAGAGTGGTGGCTCATGCTGTAAATCACAACACCTTGAGAGCCGAGGCGGTAGGTCA 2304
QY 1920 CTTGAGTCAAGGAGTTCAAAACCAAGCTGCGCCCAACATGGTGAACCCCGTCTCTACTAAA 1979
Db 2303 CTTGAGTCAAGGAGTTT-AAAACCAAGCTGACCAATGGTGAACCTTGCCTCTACTAAA 2245
QY 1980 AATCAAAAAATAGTAAAGTGTGGGCGCATGCTGTAAATCCAGCTACTTTGGAGGG 2039
Db 2244 AATAC-AAAAATCAGCAGGCTGAGTGGTGCATGCTGTAGACCCAGCTACTCCGAGGC 2186
QY 2040 TGAGGAGAGAGATTTCTTGAACCCGGGAGGAGAGTTCAGTGAAGCAGATCACAC 2099
Db 2185 TGAGGAGAGAGATTTCTTGAACCCAGGAGTGAAGTTCAGTGAAGCAGATCACAC 2126
QY 2100 ACTGCATCTCAGCTCGGGGAGAGAGAGTTCCTCTCAAAAAACCAAAAAACAAAG 2159
Db 2125 ATTGCACT-CAGCTCGGCAACAGAGAGAGTTCATCTCAAAAAACCAAAAAACAAAG 2067
QY 2160 AATTAAGCAAA 2170
Db 2066 AAAAAAGAAA 2056

```

```

RESULT 7
BI827092
LOCUS 754 bp mRNA linear EST 04-OCT-2001
DEFINITION 603075763F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:516795 5',
mRNA sequence.
ACCESSION BI827092
VERSION BI827092.1 GI:15938629
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 754)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAW1416 row: n column: 16
High quality sequence stop: 741.
Location/Qualifiers

```

```

source 1. 754
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:516795"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: BcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

```

ORIGIN

```

Query Match 7.3%; Score 286; DB 12; Length 754;
Best Local Similarity 99.7%; Pred. No. 2.7e-34;
Matches 297; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 3613 GGCAAGCAGAGTGGCAACCAATTCCTTCGGCCATGACGAGCCGAGTTTACAGAAGC 3672
Db 1 GGCAAGCAGAGTGGCAACCAATTCCTTCGGCCATGACGAGCCGAGTTTACAGAAGC 60
QY 3673 CTCATTAGCATTTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3732
Db 61 CTCATTAGCATTTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 119
QY 3733 GTCGCGAGCATCCCGGCGCCCTGCTGCGTCCGCGAGGAGGAGGAGGAGGAGGAGG 3792
Db 120 GTCGCGAGCATCCCGGCGCCCTGCTGCGTCCGCGAGGAGGAGGAGGAGGAGGAGG 179
QY 3793 CCTCCCGCCCTTACCTCCACGCGGAGCGCGCGCCAGTCAACTCTCGCACTTTGCC 3852
Db 180 CCTCCCGCCCTTACCTCCACGCGGAGCGCGCGGAGCGCGCGCCAGTCAACTCTCG 239
QY 3853 CTTGCTTGGCAGCGGATAAAGGGGGCTGAGAAATACCGGACAGGTCACCGTTGTC 3910
Db 240 CTTGCTTGGCAGCGGATAAAGGGGGCTGAGAAATACCGGACAGGTCACCGTTGTC 297

```

RESULT 8

```

BC038630
LOCUS 2772 bp mRNA linear HTC 04-MAR-2003
DEFINITION Homo sapiens, similar to hypothetical protein FLJ20489, clone
IMAGE:5263792, mRNA.
ACCESSION BC038630
VERSION BC038630.1 GI:24116283
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2772)
AUTHORS Strausberg, R.
TITLES Direct Submission
JOURNAL Submitted (15-OCT-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbio.org

```


Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAX Plate: 73 Row: 1 Column: 4

This clone has the following problem: retained intron.

Location/Qualifiers

1..2772

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_image="5263792"

/tissue_type="Brain, hippocampus"

/clone_lib="NIH MGC_95"

/lab_host="DH10B"

/note="Vector: pBluescript"

ORIGIN

Query Match 7.2%; Score 283.4; DB 11; Length 2772;

Best Local Similarity 66.1%; Pred. No. 3.1e-34;

Matches 486; Conservative 0; Mismatches 216; Indels 33; Gaps 4;

1441 TTTTCTTTTCTTTTGTGAGAGAGCTTCTGCTTTTCACTCCCTGGCTGAGGGCAGTGTG 1500

Db TTTCTTTTCTTTTGTGAGATGGAGTCTTCTCTGTTGCCAGGCT-GAGTGCATGTGTG 2129

1501 TGATCTAGCTCACTGCAACTTTCCTCCCGGTTCAAGCAATCTCTGCTTCAGGCT 1560

Db TGATCTCGGCTCACTGCACTTCTACCTCTCTGCTGGTTCAGCAATCTCTGCTTCAGGCT 2189

1561 CCCAAGTAGCTGGAGTCACTGAGCACTTCCCAACGACCCAGCTGATTTTGTATTTTGTAG 1620

Db CCGAGTAGCTGGGTCTACAGCATGACCACTAGCCCTGCTGATTTTGTATTTTGTAG 2249

1621 TAGAGATGGGATTTACATTTGTCGCAAGCTGCTCAAACTTTTCTGTCATATTTG 1680

Db TAGAGGCAAGGTTTACCATGTTGTCAGGCTGTTTGAACCTCTGACCTCAGGTGACCT 2309

1681 TTGTAACATATTGTTCTTTTCTGAGGTAGGCCCCCAGACCAAAAAATAATCTTAG 1740

Db ACCCGCTCGGCTCTCAAGTGTGTGATTACAGGATGAGCCACTCTCTGGCATAT 2369

1741 AATCCAAATCAGTGTGTGTTGTTGACCACTGCTACTTGAGAACCAACAGTGTGACAGGCG 1800

Db AACACTTTTTCAGTAAAT-----CATCTTCTATTAAATTTAGCCAGATTCCTCTA 2422

1801 CTCAGGATGAGGTGATCTCTGCTCGAAGAGAAATAGATGAATATTTCTCGGGCC 1860

Db ACAAGCAGATGATTTGATATATCAAGAA-----TCTAGCT 2458

1861 AGCGTGTGCTCATGCTCTAATCCAGACATTTGGAGGCCAAGCATGTGGATCAC 1920

Db GGGAGCTGTGCTCATGCTGTATTCACACCTTTGGAGGCTGAGCGAGTGGATCAC 2518

1921 CTGAGGTCAGAGTTCAAAACAGCTGTGCAACATGTTGAAACCCCGTCTCTACTAAAA 1980

Db CTGAGTTCAGAGTTCAGACCACTGCGCAACATGTTGAAACCCCAATATATAAT 2578

1981 ATACAAAAATTAGCTAAGTGTGTTGGCGCATGCTGTATCCAGCTACTTGGAGGCT 2040

Db ATAC-AGAAATTTGCTGGCGGTGTGTCGGCGCACCTGTGTGCTGCTTGGAGGCT 2637

2041 GAGCAGGAGAAATTTTGAACCCCGGAGGAGAGGTTGAGTGAAGCGAGATCACACCA 2100

Db GAGCAGGAGAAATCTGTTGAACCCCGGAGGAGTGTGAGCGAGCGGAGATGCGACCA 2697

2101 CTGCACTTCAGCTGGGGGAGAGCGAGATCTCTCTCAAAAAACAAAAACAAAGA 2160

Db TTGCACTTCAGCTGGGGGAGAGCGAGATCTCTCTCAAAAAACAAAAACAAAGA 2757

2161 ATTAAGCAAAATAGA 2175

Db 2758 AAAAAAAAAAAAAA 2772

RESULT 9
BC035510/c
LOCUS

DEFINITION
Homo sapiens sialophorin (gp115, leukosialin, CD43), mRNA (cdna

clone IMAGE:5094058), with apparent retained intron.
BC035510
VERSION
KEYWORDS

SOURCE
Homo sapiens (human)
ORGANISM

REFERENCE
AUTHORS
1 (bases 1 to 2076)
Strausberg, R.L., Feingold, Z.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheerz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.K., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Wozny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257
12477932
2 (bases 1 to 2076)
Strausberg, R.

Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-x@mail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
web site: <http://www.nisc.nih.gov/>

Contact: nisc.mc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAX Plate: 39 Row: P Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein

This clone has the following problem: retained intron.

FEATURES

source Location/Qualifiers
 1. .2076
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5094058"
 /tissue type="Lymph, Burkitt lymphoma"
 /clone lib="NIH_MGC 8"
 /lab host="DH10B-R"
 /note="Vector: pOTB7"

ORIGIN

Query Match 7.0%; Score 273.4; DB 11; Length 2076;
 Best Local Similarity 67.4%; Pred. No. 1.3e-32;
 Matches 487; Conservative 0; Mismatches 206; Indels 30; Gaps 6;
 QY 1444 TTTTCTTTTCTTTTTCAGAAAGAGCTTCTCTT-TCACCTGGCTGGAGGCGAGTGGTG 1502
 Db 1446 TTTTCTTTTCTTTTTCAGATGGAGTCTTTTACTTGTGCGCCAGGCTGGAATGCAATGGCGTG 1387
 QY 1503 ATCTCAGCTCACTGCAACTTTGCGCTCCCGGTTTCAAGCAATTCCTCTGCTTCAGCTCC 1562
 Db 1386 ATCTCGCTCAGTCGACCTCCACTCTCTGGTTCAAGGATCTCTCTGCTCAGCTCC 1327
 QY 1563 CAAGTAGCTGGACTACAGGACCTTCCACCATGCGCCAGCTGATTTTGTATTTTAGTA 1622
 Db 1326 CAAGTAGCTGGATCACAGTGGCGCCACACGCGCCAGCTAATTTTGTATTT- 1271
 QY 1623 GAGATGGATTTTCACTTTTGTGGCAAGCTGGTCTCAAACTTTTGTGTCFATAATGTT 1682
 Db 1270 GAAACAGGGTTTCAACATGTGTGTCAGGCTGGTCTTGAAC-CTGACCTCAG 1219
 QY 1683 GTAACATATGTTCTTCTTTTGTGAGTAGGAGGCGCCCGCCAGACCAAAATAATCTTAGAA 1742
 Db 1218 GTGATCCACTGCTCGGCTCCAAAGTCTGGGATTAAGAGCTTGAGCCACACGTC 1159
 QY 1743 TCCAAATCAGTGTGTTGTTGACCACTGTCACTTCAGAACCAAGTGTGACCAAGGCGCT 1802
 Db 1158 TGGCTGTCTATCTCTCTTTTGTGTTTCTTTTGTAGTCTTAAATATCCAGTCTCT 1099
 QY 1803 CAGGAGTAGAGTGATCTGTGCTGAAAGAGAAATAGAAATATTCCTCGGCCAG 1862
 Db 1098 TTTGCTTTGCTTTTCTCTC-AGAAATTTAAAGACATCTATTGTCG 1052
 QY 1863 GCGTGTGGCTATGCTGTAATCCAGCACTTTGGAGGCCAAGCAGTGTGATCACT 1922
 Db 1051 GCACGTGGCTCAGCATGTAATCCAGCACTTTGGAGCGCCAGGATACAGGATCAGCA 992
 QY 1923 GAGGTGAGGTTCAAAACAGCTGGCCAAACATGTTGAAACCCGCTCTCTACTAAAAAT 1982
 Db 991 GAGGTGAGGTTCAAGACAGGCTGACCAACATGTTGAAACCCCATCTCTACTAAAAAT 932
 QY 1983 ACAAAAAATTAGCTAAGTGTGGCGCATGCTGTAATCCAGCTACTTGGAGGGTGA 2042
 Db 931 AC-AAAAATATCCGGTTTGGTGGTGAAGCTGTAATCCAGCTACTTGGAGGGCTGA 874
 QY 2043 GGCAGAGAAATTTCTTGAACCCGGAGGAGAGTTGAGTGAAGGAGATCACACCACT 2102
 Db 873 GGCAGAGAAATTTTGAACCCAGGAGGCGAGGTTCTGCTGAGCCAGATCAGGCCATT 814
 QY 2103 GCATCCAGCTGGGGAGA--GAGCGAGCTTCCTCTCAAAAAACAAAAACAAAAAGA 2160
 Db 813 GCATCCAGCTGGGTAACAGAGTGAATACTCCATCTCAAAAAACAAAAACAAATCT 754
 QY 2161 ATT 2163
 Db 753 ATT 751

RESULT 10

LOCUS BUI75345 749 bp mRNA linear EST 04-SEP-2002
 DEFINITION AGENCOURT_7936105 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6011997

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

BUI75345
 BUI75345.1 GI:22689329
 EST.
 Homo sapiens (human)

ORGANISM

Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 749)
 NIH-MGC http://mgs.nci.nih.gov/.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE

Unpublished (1999)

JOURNAL

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-femail.nih.gov

COMMENT

Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM13202 row: i column: 22
 High quality sequence stop: 530.

FEATURES

Location/Qualifiers
 1. .749

source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6011997"
 /tissue type="embryonal carcinoma, cell line"
 /lab host="DH10B (phage-resistant)"
 /clone lib="NIH_MGC 92"
 /note="Organ: Testis; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 6.9%; Score 270; DB 13; Length 749;
 Best Local Similarity 64.5%; Pred. No. 7.7e-32;
 Matches 462; Conservative 0; Mismatches 230; Indels 24; Gaps 3;
 QY 1443 TTTTCTTTTCTTTTTCAGAAAGAGTCTTCTCTTTCACCTGCTGGAGGCGAGTGGTG 1502
 Db 52 TTTTCTTTTCTTTTTCAGAAAGAGTCTCTCTCTGTCAACCAGGCTGGAGTCTAGTGGTCC 111
 QY 1503 ATCTCAGCTCACTGCAACTTTTCGCTCCGGGTTTCAAGCAATTCCTCTCTCAGCTCC 1562
 Db 112 ATCTTGGCTCACTGCAACTCTGCTCTCTGGTTTCAAGTATCTCCCATCTAGCTCC 171
 QY 1563 CAAGTAGCTGGAGTACAGGCACTTCCACCATGCCCCAGCTGATTTTGTATTTTAGTA 1622
 Db 172 TCAGTAGCTGGAGCCACAGGGGTGTGCCACACATCCAGCTAATTTTGTATTTTGTG 231
 QY 1623 GAGATGGAGTTTCACTTTTGTGGCCAAAGTGTCTCAACCTTTTGTCTCATATGTT 1682
 Db 232 GAGAGGGGTTTCATCATGTGTGGCCAGGCTGTGTGCGAACTCTCTGAC- 278
 QY 1683 GTAACTATTGTTCTTTTGTCTGAGTAGGGCCCCCAGACCAAAAAATAATCTTAGAA 1742
 Db 279 -----CTAAGCGATCTGTCTGCTCGGCTCCCAAGTGTGGATAGTCTGGGAT 332
 QY 1743 TCCAAATCAGTGTGTTGTTTGAACATGTCACTTGTGAGAACACAGTGTGACAGGGCT 1802
 Db 333 TGCAGATGTGAGCTCTGACCCAGCTATCCCTTTTAA-----AGATGGGGAATAATGAA 389
 QY 1803 CAGGAGTAGAGTGTATCTCTCTCGAAGAGAAATAGATGAAATATTTCTCCGGGCCAG 1862
 Db 390 ATTTGCTTTTATGATTATACATATTATGAAGCATATTAGAAATGATTTTCTGGGCGG 449
 QY 1863 CGGTGGTGGCTCATGCGCTGTAATCCAGCACTTTGGGAGGCCAAGGCAATGTGGATCACT 1922

Db	450	GCCTGGTGGCTCAATGCTCTGTAATTTCTGGCACTTTGGGAGGCCAAGGAGGAGGAGATCACTG	509
Qy	1923	GAGTCTCAGGAGTTCACAAACACGAGCTGGCCAAACATGGTGAACACCCGCTCTCTACTAAAAAT	1982
Db	510	GAGCCAGGAGATCAAGACAGCTTGGCAACATAGCAAAACCTTGTGCTC-CAAAAAAT	567
Qy	1983	ACAAAAATTAGCTPAAGTGTGGTGGCGATGCTCTGTAATCCCACTACTTTGGGAGGGTGA	2042
Db	568	ACAAAAATTAGCTAGATGTGTGTGTGGACCTGCACTCTCAGTCTCAGTACTCAGGAGTCTGA	627
Qy	2043	GGCAGGAGAAATTTCTGAAACCCGGAGGAGAGTTCAGTGAAGCGAGATCAACCACT	2102
Db	528	GTGGGAGGATGCTTGGAGTCCAGAGATTTGAGGGTATAATGAATCTATGGTGGTGCACAACT	687
Qy	2103	GCATCCAGCCTGGGGGAGAGAGCAGATTCCTCTCAAAAAAACAAAAACAAAA	2158
Db	688	GTACTCTAGTGGGACACACAGAGGAGACCTGTCTCCAAAAAATAAAAAA	743
RESULT 11			
LOCUS	BC035771	1910 bp	linear
DEFINITION	Homo sapiens RAD1 homolog (S. pombe), mRNA (cdna clone IMAGE:5582524), with apparent retained intron.		HTC 19-NOV-2003
ACCESSION	BC035771		
VERSION	BC035771.1	GI:23958889	
KEYWORDS	HTC		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1910)		
	Klausner, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G., Strausberg, R.L., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, L.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullay, S.J., Bosak, S.A., McEwan, P.J., McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wille, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzyzanski, M.I., Skalek, U., Smalley, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.		
	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
TITLE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
JOURNAL	22388257		
MEDLINE	12477932		
PUBMED	2 (bases 1 to 1910)		
REFERENCE	Strausberg, R.		
AUTHORS	Direct Submission		
TITLE	Submitted (31-JUL-2002) National Institutes of Health, Mammalian		
JOURNAL	Gene Collection (WGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk		
	Email: cgabs-r@mail.nih.gov		
	Tissue Procurement: ATCC		
	cdna Library Preparation: Life Technologies, Inc.		
	cdna Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)		
	DNA Sequencing by: National Institutes of Health Intramural		
	Sequencing Center (NISC),		
	Gaithersburg, Maryland;		
	Web site: http://www.nisc.nih.gov/		
	Contact: nisc_mgc@hghri.nih.gov		

```

||||| 1033 CTGAGGCAAGAGATCACTTGAACCTGGAAGCGGGGTTCAGTGAGCCGAGATTGCAC 1092
|||||
QY 2099 CACTGCACTCCAGCTGGGGGAGAGAGCGAGACTTCCCTCTCAAAAAACAAAAACAAAA 2158
|||||
Db 1093 CATTGCACTCCAGCTGGGTACAGAGCAAGACTCCGCTCTCAAAAAACAAAAAATTAAA 1152
|||||
QY 2159 GAATTA 2165
|||||
Db 1153 ATAAAA 1159
|||||

RESULT 12
AL580585 929 bp mRNA linear EST 01-JUN-2003
LOCUS Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DJ015YK20 3-PRIME, mRNA sequence.
ACCESSION AL580585
VERSION AL580585.2 GI:31318853
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:12946748.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8901.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ015BF10NP1&cluster=8901.f. Contact :
Peng Liang Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ015BF10NP1.
Location/Qualifiers
1..929
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015YK20"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 6.6%; Score 259.4; DB 9; Length 929;
Best Local Similarity 62.2%; Pred. No. 2.9e-30;
Matches 507; Conservative 14; Mismatches 245; Indels 49; Gaps 6;

QY 1441 TTTTTCCTTTTTCCTGCAAGAGCTGCTCTTCACCCCTGGCTGGAGGCGAGTGGTG 1500
Db 116 TTTTTCCTTTTTCCTGCAAGAGCTGCTCTTCACCCCTGGCTGGAGGCGAGTGGTG 175
QY 1501 TGATCTCAGCTCACTGCAACTTTCGCTCCCGGGTTCAAGCAATTCCTGCTTCAGCT 1560
Db 176 TGATTCAGCTCGTGCACCTCTGCTCCCGTGTTCAGCAATTCCTGCTTCAGCT 235
QY 1561 CCCAGTAGCTGGGACTACAGGCACTTCCCAATGCCAGCTGATTTTGTATTTTAT 1620
Db 236 CCCAGTAGCTGGGACTACAGGCGGTGTCACACACTCAGCTAATTTTGTATTTCCAG 295

```

```

QY 1621 TAGAGTGGGATTTTCACTTTTGTGGCCCAAGCTGCTCACAACCTTTTTCCTGCTCATATTG 1680
Db 296 TAGAGACTG--TTTCACTATTTTGTGAGGCTGCTGTGAACTCCTGACTTCGTGA----- 349
QY 1681 TTGTAATAATTTGTTTCTGAGGTAGGCGCCCGCCAGACCAAAAAAATAAATCTTAG 1740
Db 350 -----TCGCGCCCACTCAGCCTCCCAAGTGTGGATTTACAGGCGAG 392
QY 1741 AATCCAAATCAGTGTGTTGTTGACCACTGTCTACTTTGAGAACCCACAGTGTGACAGGGC 1800
Db 393 CCACC-----TTGCGCGCGCGAAGTTAAAGATGCGAGTTTGGGATAAGACC 437
QY 1801 CTCAGGAGTAGAGGTGATCTCTCTCGAAAGAGAGAAATAGAAATATTTCTCGGCGCC 1860
Db 438 ATCATAGGATATTTTCTCCCTT-----TACTCATAAAGTTGTGAGCTTTGGCC 489
QY 1861 AGCGGTGGTGGCTATGCTGCTGTAATCCAGCACTTTGGGAGGCCAAGCATGTGATCAC 1920
Db 490 GGGCGCGATGGCTCAGCCCTGTTATCTAGTACTTTGGGAGGCTGAGCGCGGAGATCAC 549
QY 1921 CTGAGGTCAGGAGTTCAAAACCCAGCCTGGCCCAACATGGTGAACCCCGCTCTTACTTAAA 1980
Db 550 CTGAGGTCAGGAGTTTGTGACCAAGCTGGCCCAACATGACAAAACCCCTGTCTTACTTAAA 609
QY 1981 ATACAAAAAATTAGCTAAGTGTGCTGGCGCATCGCTGTATATCCAGCTACTTGGGAGGGT 2040
Db 610 ATAC--AAAACTAGCCAGGTGTGCTGGCGCACCTGCAGTCCCACTATTCGGGAGGCT 667
QY 2041 GAGGCGAGGGAATTTCTTGAACCCGGAGGAGAGGTTGCGAGTGAAGCGAGATCACACCA 2100
Db 668 GAGGCGAGGAGAAATCACCCGAAACCTGGGAGGCGAAGGTTGTAGCGAGCCGAGATCTTGCCA 727
QY 2101 CTGCACTCCAGCTGGGCGAGAGAGCGAGCACTTCCTCTCAAAAAACAAAAACAAAAAGA 2160
Db 728 TCACACTCCAGCTGGAGGACAGAGCAAACTGTGTCACAAAAA--AAAAAAGT 786
QY 2161 ATTAAGCAAAATTAGCATTCAGAGAGAACCTGAAGGGGCTCAGACCACTACAGATTTC 2220
Db 787 TGTAGCTTAAATAATCCCTTCACGAGAGCTTCTTTTGTCTTAAAMCTTTCTCTTTT 846
QY 2221 TGTGCCACATGCCAAGTACTTCTGAGGCGATGACTG 2255
Db 847 TTTAAWCTGTYTGCCKGKTTTGAGGCGGAAACAG 881

```

```

RESULT 13
BU616025 693 bp mRNA linear EST 23-SEP-2002
LOCUS UI-H-DF0-ben-m-20-0-UI.s1 NCI CGAP DF0 Homo sapiens cDNA clone
DEFINITION UI-H-DF0-ben-m-20-0-UI 3', mRNA sequence.
ACCESSION BU616025
VERSION BU616025.1 GI:23282240
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 693)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone Distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 9-305, >ALU (matched complement) 326-387, >L1MB7#LINE/L1
(matched complement) 388-676, >ALU

```

Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES source

```

FEATURES
source
Location/Qualifiers
1. .693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DF0-ben-m-20-0-UT"
/tissue_type="Subchondral Bone"
/dev_stage="Adult"
/lab_host="DHI08 (Life Technologies)"
/clone_lib="NCI CGAP DF0"
/note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP DF0 is a cDNA library containing the following
tissue(s): Subchondral Bone. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed,
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GTTAAAGCGTC.
TAG TISSUE=subchondral bone
TAG LIB=UI-H-DF0
TAG SEQ=GTTAAAGCGTC"

```

ORIGIN

	Query Match	6.6%	Score 259.2;	DB 13;	Length 693;
	Best Local Similarity	64.7%;	Pred. No. 3.7e-30;		
	Matches 465;	Conservative 0;	Mismatches 216;	Indels 38;	Gaps 4;
QY	1441	TTTTTTTTTTTTTTTGAGAGAGCTTGCTCTTACCCCTGGCTGAGAGGAGAGTG			1500
Dbb	1	TTTTTTTTTTTTTTTGAGACAGATTTGCTCTATCACCCAGCTGGAGTCAGTGCA			60

QY
db

1441 TTTT...TTTGTGAAACAGCTCTTGCTCTTCACCCCTGGCTGGAGGGCAGTGGTG 1500
1 TTTT...TTTGTGACACAGTTTTTGCTCTATCACCCAGCTGGAGTGCAGTGGCA 60

1501 TGATCTCAGCTCACTGCAACTTTGGCTCCGGGTTCAAGCAATTCTCTGCTTACGCT 1560
61 CAATCACACTTCACAGCAGCCTGTAACTTTGGGCTCAAGCGGTCTCTCCACCTTACGCT 120

1561 CCACAAGTAGCTGGGACTACAGGCACCTCCCAACATGCCCAGCTGATTTT---GTATTTT 1617

121 CCCGGSTAACTGGAACTCGAGGTGTGTGCACATATGCCCAGCTAAATTTTTAAAAAATAAT 180

1618 TAGTAGAGATGGGATTTCACTTTTGTTGGCCAGCTGGTCTCAAACTTTTGTGTGCATAA 1677
181 TTATAGAGATAGAGTCTCTCTATGTTACTTAGCTGATCTCAAACTCTGG----- 231

1678 TTGTTGTAACATATTGTTCTCTTTTGCTGAGGTAGGGCCCCCAGACCAAAAAATAAATCT 1737
232 -----GCTCAGGCAGTCTCTCTGTTNTCCGCCCTCCCAAGTGTGGGATTAAATAG 281

1738 TAGAATCCAAATCAGTGTGTGGTTTGACCACTGCATTGAGAACCAACAGTGTGACCAG 1797
282 GGGTGTGAGGCCACCAACCCAGCCCAATACTGTTTTTAAAAAGACTTAAT^a----- 333

1798 GGCCTCAGGAGTAGAGGTGATCTCTGCTCGAAGAGAAATAGATGAAAAATATTCTCCGG 1857

334 -----TGAGACATCTTCACATACCATCAATTCACATATTTAAAGTGTAACAATCAG 385

1858 GCACGGCGTGTGGTCTATGCCGTGTAATCCACGACATTTGGAGAGCCCAAGGCATGTGGAT 1917
387 TCCGGGCATGTGGTCTATGCGTGTGTAATCCACGACATTTGGAGAGCCCAAGGCATGTGGAT 445

1918 CACCTGAGGTGAGGAGTTCAAAACACGCGCTGSCCAACATGTGTGAACCCCGTCTCTACTA 1977
|||||
447 CACCTGAGCGCGGAGTTTCAGACGCGCGCTGSCCAACATGTGTGAACCCCGTCTCTACTA 1977
|||||

[illegible]

... TTTT CCGGCTGGAGTCGGGTGATCCCGTGATCCCCAGTTACTCAGGAG 565

QY	2038	GTTGAGCAGGAGAAATTTCTTGAAACCGGAGGAGGAGGTTTGAGTGAAACCGGAGATCACA	2097
Db	566	GCTGAGCAGGAGAAATCTTTTGACCTGGGAGGCGGAGTTTGAGTGAGCTGAGATCACA	625
QY	2098	CCA CTGCACTCCAGCCTGGGGAGAGGAGAGATCTCCTCTCAAAAAA CAAAAACAA	2156
Db	626	CCACTTACTCAAGCCTGGGTGACACAGCGAGATCTTGTTCTCAAAAAA CAAAAA	684

RESULT 14
AF101960

LOCUS	AF101960	Human Homo sapiens	2971 bp	DNA	linear	GSS 06-NOV-2000
DEFINITION	AF101960	Human Homo sapiens	genomic clone	ptWB59.14,	genomic survey	
ACCESSION	AF101960		sequence.			
VERSION	AF101960					
KEYWORDS	AF101960.1	GI:4193786				
SOURCE	GSS.					
ORGANISM	Homo sapiens	(human)				
	Homo sapiens					
	Eukaryota; Metazoa;	Chordata; Craniata;	Vertebrata; Euteleostomi;			
	Mammalia; Eutheria;	Primates; Catarrhini;	Hominidae; Homo.			
REFERENCE	1 (bases 1 to 2971)					
AUTHORS	Bepko,G., O'Brian,K.C., Kim,Y.C.,	Schreiber,G. and	Pitterle,D.M.			
TITLE	A 1.4-Mb high-resolution physical map and contig of chromosome segment 11p15.5 and genes in the LOH1A metastasis suppressor region					

JOURNAL
MEDICAL

99134294
MEDLINE
PUBMED
COMMENT
Contact: Repler G
Medicine and Radiology
Duke University Medical Center
Box 2610, MSRB, Room 117, Durham, NC 27710, USA
Part of a 1.4 megabase contig including the LOH1A metastasis
suppressor region Bin T
Class: unknown.

FEATURES	SOURCE
----------	--------

```
/organism="Homo sapiens"  
/mol_type="genomic DNA"  
/db_xref="taxon:9606"  
/map="llp15.5"  
/clone="pTWB59.14"  
/clone_lib="Human"
```

ORIGIN

Query Match	5.5%;	Score 255.8;	DB 28;	Length 2971;
Best Local Similarity	65.4%;	Pred. No. 5.1e-30;		
Matches 478;	Conservative 0;	Mismatches 177;	Indels 76;	Gaps 4;
1442	TTTTTTTTTTTTTTTGAGAGAGTCTTGCTCTCTTACCCCTGGCTGGAGGCGAGTGTGT	1501		
478	TTTATTTATTTTTTCCGAGACAGTGTCTCACTCTGTGTGCCAGACTGGAGTGCAGTGGCAC	537		
1502	GATCTCAGCTCACTGCAACTTTTGCCCTCCGGGTTCAAGCAATTCCTCTCCTCAGCGCTC	1561		
538	AACTCTGGCTCACTGCAACTCTCGCTCTCGGGTTCAAGCGATTCTCTCGCTCAGCGCTC	597		
1562	CCAAGTAGCTGGGACTACAGGCACCTTCCCACTATGCCAGCTCATTTTGTATTTTGTAGT	1621		
598	CCAAGTAGCTGGGATTACAGGCGCCGCCACACACAGCTAATTTTGTATTTTGTAGT	657		
1622	AGAGATGGGATTTCACTTTTGTGGCCAAAGCTGGTCTCTCAAACTTTTTTGTCTCATAAATTGT	1681		
658	AGAGATGAGGTTTCAAAATGTTGGCCAGGCTGTCTCAAACTCCT-----	702		
1682	TGTAACTATTGTTCCTTTTGTCTGAGTGAAGGCCCCCGAGACCAAAAAATAAATCTTAGA	1741		
703	-----GACCTGTGTATCCACCCACCTCAGCCCTCCCA	733		
1742	ATCCAAATCAGTGTGTGGTTTGACCACTGTCTACTTGTAGAACCAACAGTGTGACAGGGCC	1801		

Db 734 AAGTGTGGGATTACAGGATTAAGCCACCCGATCTGTGGCCCAACATTATAT----- 783

QY 1802 TCAGGAGTAGAGGTATCTCTGCTCGAAAGAGAGAAATAGAAATATTTCTCCGGGCCA 1861

Db 784 -----TTTATGGTGAAGACTGAATGTTTCTTAAGATCAGAAACAGCCA 827

QY 1862 GCGGTGGTGGCTCATGCTCTATATCCAGCACATTTGGAGGCCAAGCATGTGGATCACC 1921

Db 828 GGCACAGTGGCTCAGCGCTTAATCTCTGGCACATTTGGAGGCCGAGATCAC- 886

QY 1922 TGAGGTGAGGAGTTCACAAACAGCGTGGCCAAATGTTGAACCCCGTCTCTACTAAAA 1981

Db 887 -GAGTTCAGAGATTCAGACCATCTCTGGCTTAACACCGTGAACCCCGTCTCTACTAAAA 945

QY 1982 T-----ACAAAAATTAGCTAAGTGTGGTGGCGCATGCTCTTAATCCAGCATCTGGGAG 2037

Db 946 TAAAAATAAAAAATTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1005

QY 2038 GGTGAGGAGGAGAAATTTCTTGAACCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2097

Db 1006 GCTGAGGCGGGAGATGGTGTGAACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1065

QY 2098 CCACTGACCTCCAGCTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2157

Db 1066 CCACTGACCTCCAGCTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1125

QY 2158 AGAATTAAAGCA 2168

Db 1126 GAAATGAACA 1136

RESULT 15

AL578531/c

LOCUS

DEFINITION

AL578531 Homo sapiens HELA CELLS COT 25-NORMALIZED EST 01-JUN-2003

ACCESSION

AL578531

VERSION

AL578531.2

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

AUTHORS

Full-length cDNA libraries and normalization

TITLE

Unpublished (2001)

JOURNAL

On Feb 16, 2001 this sequence version replaced gi:12942694.

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 2057.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DK001AB07NP1&cluster=2057.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DK001AB07NP1.

Location/Qualifiers

1. 958

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DK001YC13"

/cell_type="HELA CELLS COT 25-NORMALIZED"

/cell_line="HELA"

/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES

SOURCE

ORIGIN

Query Match 6.5%; Score 255.2; DB 9; Length 958;

Best Local Similarity 62.2%; Pred. No. 1.2e-29;

Matches 463; Conservative 23; Mismatches 243; Indels 15; Gaps 6;

QY 1441 TTTTCTTTTCTTTTGTGAGAAAGAGTCTGCTCTTTTCCACCTGGCTGGAGGAGGAGTGGTG 1500

Db 940 TTTTCTTTTCTTTTGTGAGACAGRGCCCTCTTTTGTGCMCAGRCCTGMAATGAATGGGG 881

QY 1501 TGATCTCAGCTCAGTCCAACTTTTGGCTCCCGGTTTCAAGCAATCTCTCTGCTTCAAGCT 1560

Db 880 CGWCTCTCAGCTWACTGCAGCTCCMCCHACTGGGTTAAAGTSTCTCCCACTCCGCT 821

QY 1561 CCCAAGTAGCTGGGACTCAGGCACTTCCACAC---ATGCCAGCTGATTTTGTATTTT 1617

Db 820 CCCGAGWAGCCGGGACCAAGGCGTGTATMACTGTACGCTGGCTAAATTTTGTATTTT 761

QY 1618 TAGTAGAGATGGGATTTCACTTTTGGCCAGCTGGTCTCAAACTTTTGTCTGTCATTA 1677

Db 760 TATGGAGACGAG-TTTCMCATGTTGCCAGATGKTCTGGMACTCCT---GGGCTMAA 705

QY 1678 TTGTTGTTAACTATTTGTTCTTTGCTGAGGTAGGGCCCCCAGACCAAAAAAATAATCT 1737

Db 704 GTGGTCTGCCACTTCCGCTCCCAAGTGTGGGATTATGGCC---ATTAGCCACATCT 848

QY 1738 TAGAATCCAAATCAGTGTGTTGGTTTGACCACTGTCACTTGAGAACCAAGTGTGACCA 1797

Db 647 GCCTGGCAATATTTTGTATACCTCTCTTATRCCAAGAAWAAGTCTTAATTTCTTT 588

QY 1798 GGCCTCAGGAGTAGAGGTGATCTCTGCTCGAAAGAGAAATAGAAATTTCTCCGG 1857

Db 587 ACTATGTGAAGTAACTTATAAGGTAAATAGAAAGCAAGCAAAAGTGAAGTTCATGCT 528

QY 1858 GCCAGGCTGTGGTGGC---TCATGCTGTAAATCCCAAGCACTTTGGAGGCCAAGGATGTG 1914

Db 527 ATAATCCCAAGCGCGCTGTAAACGCTGTGATCCCAAGCCTTTTGAGAGCTGAGCGGCTG 468

QY 1915 GATCACCTGAGTGCAGGAGTTCAAAACAGCTGGCCAACTGGTGAACCCCGTCTCTA 1974

Db 467 GATAACTTAAAGTTCAGGAGTGTAGAGACCAAGCTGACCAACATGGTGAATAACTGTCTCTA 408

QY 1975 CTAAAAATACAAAAAATTAGCTAAGTGTGGTGGCCATCGCTGTAAATCCCAAGCTACTTGG 2034

Db 407 CTAAAAAT--AAAAAATTAGCTGGGTTTGGTGGCAACGCTGTAAATCCCAAGCTACTTGA 350

QY 2035 GAGGTGAGGAGGAGAAATTTCTTGAACCCGGGAGGAGGAGGTTGCAAGTGAAGCGAGATC 2094

Db 349 GAGGTGAGGAGGAGAAATCGCTTGAACCTGGGAGGAGGAGGTTGCAAGTGAAGCGAGATG 290

QY 2095 ACACCACTGCCTCCAGCTGGGAGAGAGAGAGACTTCTCTCAAAAAACAAAAAAC 2154

Db 289 GTGCCATGCACTCGACCTGGGCAAGAGAGCGGAAACTCTCTMAAAARRDKKWA 230

QY 2155 AAAAGAAATTAAGCAAAATTAGACAT 2178

Db 229 AAGTAACAGAAAAAAGAGTGAAT 206

Search completed: May 13, 2004, 17:08:32

Job time : 14376.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:19 ; Search time 24113.2 Seconds
(without alignments)
7114.428 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3958
Perfect score: 3958
Sequence: 1 aggaattacatttaaat.....gggaactcaccgacgcggg 3958

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3958	100.0	4150	6	AX528729	AX528729 Sequence
2	3958	100.0	4150	9	AF320307	AF320307 Homo sapi
3	3928.2	99.2	159351	2	AC018783	AC018783 Homo sapi
4	3917.8	99.0	115278	9	AL139819	AL139819 Human DNA
5	1827.2	46.2	185644	2	AL391723	AL391723 Homo sapi
6	1799.8	45.5	1970	6	AX431388	AX431388 Sequence
7	1758	44.4	161827	2	AC139668	AC139668 Papio ham
8	1448.2	36.6	179879	2	AC026883	AC026883 Homo sapi
9	816.6	20.6	190784	2	AC139669	AC139669 Lemur cat
10	521	13.2	17088	4	AY241932	AY241932 Bos tauru
11	498.4	12.6	617	6	AX357255	AX357255 Sequence
12	460	11.6	164201	9	AL157935	AL157935 Human DNA
13	447.2	11.3	200288	2	AC119725	AC119725 Homo sapi
14	433	10.9	175842	9	AL731547	AL731547 Human DNA
15	423.8	10.7	190508	9	AC005037	AC005037 Homo sapi
16	415.2	10.5	44496	9	AC004760	AC004760 Homo sapi
17	415.2	10.5	190814	9	AC006111	AC006111 Homo sapi
18	408.8	10.3	183690	9	AC091180	AC091180 Homo sapi
19	407.4	10.3	170528	9	AC083866	AC083866 Homo sapi
20	407.4	10.3	217615	9	AC011005	AC011005 Homo sapi
21	404.2	10.2	137289	9	AL161909	AL161909 Human DNA
22	403	10.2	189179	2	AC090265	AC090265 Homo sapi
23	402.8	10.2	181597	9	AC090515	AC090515 Homo sapi
24	402.8	10.2	187734	9	AC025918	AC025918 Homo sapi
25	401.8	10.2	204963	2	AC145617	AC145617 Homo sapi
26	394.2	10.0	179599	2	AC13278	AC13278 Homo sapi
27	394.2	10.0	195558	9	AC007448	AC007448 Homo sapi
28	392.2	9.9	155559	2	AL133336	AL133336 Homo sapi
29	392.2	9.9	169835	9	AP005717	AP005717 Homo sapi
30	391.6	9.9	171641	9	AC018512	AC018512 Homo sapi
31	390	9.9	139632	9	AC145334	AC145334 Pan trogl
32	390	9.9	171555	2	AC087537	AC087537 Homo sapi
33	390	9.9	181891	9	AC023356	AC023356 Homo sapi
34	389.8	9.8	67199	9	AC025162	AC025162 Homo sapi
35	388.4	9.8	172421	2	AC013279	AC013279 Homo sapi
36	386.2	9.8	112364	9	AC109927	AC109927 Homo sapi
37	386	9.8	73162	9	AC105095	AC105095 Homo sapi
38	386	9.8	197092	2	AC036155	AC036155 Homo sapi
39	384.8	9.7	140529	2	AC011486	AC011486 Homo sapi
40	384	9.7	133475	9	AC006254	AC006254 Homo sapi
41	382.4	9.7	189639	9	AC010997	AC010997 Homo sapi
42	381.6	9.6	80775	9	AC068539	AC068539 Homo sapi
43	381.2	9.6	150681	9	AC006011	AC006011 Homo sapi
44	381.2	9.6	187064	9	AC011476	AC011476 Homo sapi
45	381.2	9.6	188413	2	AC145834	AC145834 Homo sapi

ALIGNMENTS

RESULT 1
AX528729
LOCUS AX528729
DEFINITION Sequence 15 from Patent WO0236780.
ACCESSION AX528729
VERSION AX528729.1 GI:25172800
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Prouty, S.M., Zhang, L. and Stenn, K.S.
TITLE Stearoyl-coa desaturase gene promoter
JOURNAL Patent: WO 0236780-A 15 10-MAY-2002;

1981	DB	ATACAAAAAATTAGCTTAAGTGTGGTGGGCATGCGCTGTAATCCCGAGCTTACTTGGGAGGGT	2041
2041	QY	GAGGCAGGAGAAATTTCTTGAACCCGGGAGGACAGAGTTGCAGTGAACGCGAGATCACACCA	2101
2041	DB	GAGGCAGGAGAAATTTCTTGAACCCGGGAGGACAGAGTTGCAGTGAACGCGAGATCACACCA	2100
2101	QY	CTGCATCTCAGGCTCTGGGGGAGAGCGGACATTCCTCTCAAAAAAACAACCAAGG	2160
2101	DB	CTGCATCTCAGGCTCTGGGGGAGAGCGGACATTCCTCTCTCAAAAAACAACCAAGG	2160
2161	QY	ATTAAAGCAATTAGACATTCGAGAGAGAACCTGAAAGGGGTGCACACGTTACAGATTTC	2220
2161	DB	ATTAAAGCAATTAGACATTCGAGAGAGAACCTGAAAGGGGTGCACACGTTACAGATTTC	2220
2221	QY	TGTGCCCATGTCACCAAGTACTTCTGAGGCAATGACTTGGATGAGCTGTCCACATCTGAAATCA	2280
2221	DB	TGTGCCCATGTCACCAAGTACTTCTGAGGCAATGACTTGGATGAGCTGTCCACATCTGAAATCA	2280
2281	QY	TCCAGTCTTGTTCGAACTTTTCACACCGGACAGGAGCCAGGACTGGAATGCAATCTCCT	2340
2281	DB	TCCAGTCTTGTTCGAACTTTTCACACCGGACAGGAGCCAGGACTGGAATGCAATCTCCT	2340
2341	QY	GGTCACTGGCCAGAGATTGGCTTGACCTTGACACCAAGTGGCCACAAGAGAGCTGCTT	2400
2341	DB	GGTCACTGGCCAGAGATTGGCTTGACCTTGACACCAAGTGGCCACAAGAGAGCTGCTT	2400
2401	QY	AGTCTACCTCCACAGGAATCCAGGTGCTGTCTTCTCTGGGAAGTGAATCATTTGGCGCAG	2460
2401	DB	AGTCTACCTCCACAGGAATCCAGGTGCTGTCTTCTCTGGGAAGTGAATCATTTGGCGCAG	2460
2461	QY	CACCTCCGTATTTTCTCTCTTCCACGGGGAAGGATCCTAGGGCAGTATTTGGGAAGACA	2520
2461	DB	CACCTCCGTATTTTCTCTCTTCCACGGGGAAGGATCCTAGGGCAGTATTTGGGAAGACA	2520
2521	QY	TGGGATGGAAGGACACCGGTGAATGCATAGCCTGCCTGGTTCTGAGCTCTCATGTTAA	2580
2521	DB	TGGGATGGAAGGACACCGGTGAATGCATAGCCTGCCTGGTTCTGAGCTCTCATGTTAA	2580
2581	QY	GGCTCTTACAGACACGGAAGAATGGGGGCACAGGACACAGATCAGTAGGCTCAGAGCATC	2640
2581	DB	GGCTCTTACAGACACGGAAGAATGGGGGCACAGGACACAGATCAGTAGGCTCAGAGCATC	2640
2641	QY	TCAGGACCGAGGGCAATATGTTCTTGACGAGGGATTAAGCTTGGGCTCTCATATGGT	2700
2641	DB	TCAGGACCGAGGGCAATATGTTCTTGACGAGGGATTAAGCTTGGGCTCTCATATGGT	2700
2701	QY	GTTTCTGGGCTCAACTGCCAGCTCCGTCATCTACTGCTGCTGTCGACCATGGGCAAGTTA	2760
2701	DB	GTTTCTGGGCTCAACTGCCAGCTCCGTCATCTACTGCTGCTGTCGACCATGGGCAAGTTA	2760
2761	QY	TTCCATCTCTCCATATCTTTCTTCTTATTCACATTAATGGAATGAGGTACCCACCTCC	2820
2761	DB	TTCCATCTCTCCATATCTTTCTTCTTATTCACATTAATGGAATGAGGTACCCACCTCC	2820
2821	QY	CAGGGTCACAGAGGCTTACAGAAAAGATTCTGTGTAATTTGGCTTGCAGTAATAATTC	2880
2821	DB	CAGGGTCACAGAGGCTTACAGAAAAGATTCTGTGTAATTTGGCTTGCAGTAATAATTC	2880
2881	QY	AATACCTTGGCAGCTATTCTTATTCACATCCAAAGCCCTTTGCGCTGCTGGGTGAAAA	2940
2881	DB	AATACCTTGGCAGCTATTCTTATTCACATCCAAAGCCCTTTGCGCTGCTGGGTGAAAA	2940
2941	QY	CACATGTCAGTGTTCCTGACCGGTTTCCAAAGAGATTTCCAAATTTACAACTTGCAG	3000
2941	DB	CACATGTCAGTGTTCCTGACCGGTTTCCAAAGAGATTTCCAAATTTACAACTTGCAG	3000
3001	QY	TCTTGAGAAATCTCCAAAAATCCCGCACATCTCTGGAGGGCGGGCTTGGGATGGGAC	3060
3001	DB	TCTTGAGAAATCTCCAAAAATCCCGCACATCTCTGGAGGGCGGGCTTGGGATGGGAC	3060
3061	QY	TGCCCCCGGGTCTCTGACAGATGCGTGGCGGACGAGCACACACAGCAGCGCTG	3120
3061	DB	TGCCCCCGGGTCTCTGACAGATGCGTGGCGGACGAGCACACACAGCAGCGCTG	3120

[illegible]

KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
		Homo sapiens (human)			
		Homo sapiens			
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
		1 (bases 1 to 4150)			
		Zhang, L., Ge, L., Tran, T., Stenn, K. and Prouty, S.M.			
		Isolation and characterization of the human stearyl-CoA desaturase			

gene promoter: requirement of a conserved CCAAT cis-element
 Biochem. J. 357 (Pt 1), 183-193 (2001)
 21308466
 MEDLINE
 11415448
 PUBMED
 2 (bases 1 to 4150)
 REFERENCE
 Zhang, L., Ge, L., Tran, T., Stenn, K. and Prouty, S. M.
 Direct Submission
 Submitted (09-NOV-2000) Skin Biology TRC, Johnson and Johnson,
 CPW, 199 Grandview Road, Skillman, NJ 08558, USA
 JOURNAL
 FEATURES
 Location/Qualifiers
 source
 1..4150
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /tissue_type="scalp skin keratinocytes; liver; hair
 follicles; brain"
 3602..>4150
 /gene="SCD"
 /gene="SCD"
 3602..3606
 /gene="SCD"
 3606..3606
 /gene="SCD"
 3869..3872
 /gene="SCD"
 3924..3927
 /gene="SCD"
 <3958..>4129
 /gene="SCD"
 /product="stearyl-CoA desaturase"
 /note="3602"
 <3958..4129
 /gene="SCD"
 /number=1
 3958..4102
 /gene="SCD"
 4103..>4129
 /gene="SCD"
 /codon_start=1
 /product="stearyl-CoA desaturase"
 /protein_id="AAK54510.1"
 /db_xref="GI:14150491"
 /translation="MPAHLQDD"
 4130..>4150
 /gene="SCD"
 /number=1
 ORIGIN
 Query Match 100.0%; Score 3958; DB 9; Length 4150;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGGAATTCATCCATTAAATCATACAAATTAATGGCTTTAGTATATTCACAGGTTGTC 60
 Db 1 AGGAATTCATCCATTAAATCATACAAATTAATGGCTTTAGTATATTCACAGGTTGTC 60
 QY 61 ATCCATCAAAATCCATTTTAAAGACAGTTTATTAATCTCCAAATAATAAACCTGCATTCCTT 120
 Db 61 ATCCATCAAAATCCATTTTAAAGACAGTTTATTAATCTCCAAATAATAAACCTGCATTCCTT 120
 QY 121 AGCCATCAACCCCAACATCCATCCATCCCTCCATCCCTCCATCCATCCATCCATCCATCCAT 180
 Db 121 AGCCATCAACCCCAACATCCATCCATCCCTCCATCCCTCCATCCATCCATCCATCCATCCAT 180
 QY 181 TTCTGCTCTCTATAAATTTGCAATTCCTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTT 240
 Db 181 TTCTGCTCTCTATAAATTTGCAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTT 240
 QY 241 GTGAGACTTTGTGACTGGCTGCTTCACTTAGCAATTTCACTTAGCAATTTCACTTAGCAATTT 300
 Db 241 GTGAGACTTTGTGACTGGCTGCTTCACTTAGCAATTTCACTTAGCAATTTCACTTAGCAATTT 300
 QY 301 CAGTACTTAGCAGTACTTCACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 360
 Db 301 CAGTACTTAGCAGTACTTCACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 360

QY 361 CCCATATCATATTTATAGACAGAGTTCTCATCTGTCAACCCAGGCTGAGTGCAGTGGC 420
 Db 361 CCCATATCATATTTATAGACAGAGTTCTCATCTGTCAACCCAGGCTGAGTGCAGTGGC 420
 QY 421 ACAATCATAGTCACTGTAACTCAAACTCTCTGGGCTCAAGTGAATCTTACTACCTCAGCC 480
 Db 421 ACAATCATAGTCACTGTAACTCAAACTCTCTGGGCTCAAGTGAATCTTACTACCTCAGCC 480
 QY 481 TCCAGAGTGTAGTACGACTACAGGACACACAGCCATACCTGCTAAATTTTTTTTAAAT 540
 Db 481 TCCAGAGTGTAGTACGACTACAGGACACACAGCCATACCTGCTAAATTTTTTTTAAAT 540
 QY 541 TTTTCATTTTATGATTTCAATTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 600
 Db 541 TTTTCATTTTATGATTTCAATTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 600
 QY 601 ACTTTGTTATCCAGGCTGGAGGCGAGTGGCATGGTGACAGCTGAGCAGCCCTTGACTTCCT 660
 Db 601 ACTTTGTTATCCAGGCTGGAGGCGAGTGGCATGGTGACAGCTGAGCAGCCCTTGACTTCCT 660
 QY 661 GGGCTCAAGTGAATCTCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGTGTCA 720
 Db 661 GGGCTCAAGTGAATCTCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGTGTCA 720
 QY 721 CCATGCTGCTGATATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 780
 Db 721 CCATGCTGCTGATATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 780
 QY 781 GTACAGTGGCGTAAATATAGTCACTGACGCTCCCTCTCTGCGCTCAAGCAATCCGCTG 840
 Db 781 GTACAGTGGCGTAAATATAGTCACTGACGCTCCCTCTCTGCGCTCAAGCAATCCGCTG 840
 QY 841 GCCTCAGCATCTGAGTAGCTGGGACTACAGGCTTGTGCCACAGCCAGCCAGCTTAAGTTT 900
 Db 841 GCCTCAGCATCTGAGTAGCTGGGACTACAGGCTTGTGCCACAGCCAGCCAGCTTAAGTTT 900
 QY 901 AAAAATGATTTTGTATAGAGGAGTCTTCTATGTTGCTCAGGCTGTATTTTATTTG 960
 Db 901 AAAAATGATTTTGTATAGAGGAGTCTTCTATGTTGCTCAGGCTGTATTTTATTTG 960
 QY 961 TTGAGACAAGGTCTCACTATGTTGCCATGATCCCCCAGCTCCCAAGTGTCTCA 1020
 Db 961 TTGAGACAAGGTCTCACTATGTTGCCATGATCCCCCAGCTCCCAAGTGTCTCA 1020
 QY 1021 TCTTATCTGTTTCACTTAGTGTGACAGCATTTAGTGTGTTTCCATTTTGAACATTA 1080
 Db 1021 TCTTATCTGTTTCACTTAGTGTGACAGCATTTAGTGTGTTTCCATTTTGAACATTA 1080
 QY 1081 TGAATAATACTCCAGTGAATTTCACTATATACATTTGTGGGCATATGTTTCAATTTCT 1140
 Db 1081 TGAATAATACTCCAGTGAATTTCACTATATACATTTGTGGGCATATGTTTCAATTTCT 1140
 QY 1141 GTTGGTTTATATCTAGGAGTGAATTTCTGGATCCCGGGTAAATTTTGAACAGCAG 1200
 Db 1141 GTTGGTTTATATCTAGGAGTGAATTTCTGGATCCCGGGTAAATTTTGAACAGCAG 1200
 QY 1201 TTCAGGGGAAGAAAACCTTGGGAAAATGAAGCATGTTTAGAAAATCAGCAGTGCAGGG 1260
 Db 1201 TTCAGGGGAAGAAAACCTTGGGAAAATGAAGCATGTTTAGAAAATCAGCAGTGCAGGG 1260
 QY 1261 GTTTTTCGGAGTTTATTTTATATTTCTGTTGACAAATGTGCAGTTTGTAGATACAG 1320
 Db 1261 GTTTTTCGGAGTTTATTTTATATTTCTGTTGACAAATGTGCAGTTTGTAGATACAG 1320
 QY 1321 TTATCTAGTGTGAGAGTGAATTAAGCTGGAAATAGGGCGTTTCAAGTAAAAATCATGA 1380
 Db 1321 TTATCTAGTGTGAGAGTGAATTAAGCTGGAAATAGGGCGTTTCAAGTAAAAATCATGA 1380
 QY 1381 AGCACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATAATAAAAAATCAAA 1440
 Db 1381 AGCACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATAATAAAAAATCAAA 1440

1441 TTTTTTTTTTTTTTTGAGAAAGAGTCTTGCTCTTTTCAACCTGGCTGAGGGCAGTGGTG 1500
1441 TTTTTTTTTTTTTTTGAGAAAGAGTCTTGCTCTTTTCAACCTGGCTGAGGGCAGTGGTG 1500
1501 TGNATCTCAGCTCACTGCACTTTTCGGCTCCCGGGTTCAAGCAATTCCTCGCTTCAGCCT 1560
1501 TGNATCTCAGCTCACTGCACTTTTCGGCTCCCGGGTTCAAGCAATTCCTCGCTTCAGCCT 1560
1561 CCCAAGTAGCTGGCACTACAGGCACATCCACCAATGCCAGCTGATTTTGTATTTTTAG 1620
1561 CCCAAGTAGCTGGCACTACAGGCACATCCACCAATGCCAGCTGATTTTGTATTTTTAG 1620
1621 TAGAGATGGGATTTCACTTTTGTGGCAAGCTGGTCTCAAACTTTTGTCTCATATAATTG 1680
1621 TAGAGATGGGATTTCACTTTTGTGGCAAGCTGGTCTCAAACTTTTGTCTCATATAATTG 1680
1681 TTGTAACTATTTGTTCTTTTGTGGTGGGCCCCCAGACCAAAAAATAAATCTTAG 1740
1681 TTGTAACTATTTGTTCTTTTGTGGTGGGCCCCCAGACCAAAAAATAAATCTTAG 1740
1741 AATCCAAATCAGTGTGTTGGTTTGACCACTGTCACTTTGAGAACCAAGTGTGACCAAGGC 1800
1741 AATCCAAATCAGTGTGTTGGTTTGACCACTGTCACTTTGAGAACCAAGTGTGACCAAGGC 1800
1801 CTCAGAGTAGAGTGATCTCTGCTCGAAAGAGAAATAGAATGAAATATTTCTCGGGGCC 1860
1801 CTCAGAGTAGAGTGATCTCTGCTCGAAAGAGAAATAGAATGAAATATTTCTCGGGGCC 1860
1861 AGGCGTGGTGGTCACTGCTGTAATCCAGCACTTTGGAGGCCAAGGATGTGATCAC 1920
1861 AGGCGTGGTGGTCACTGCTGTAATCCAGCACTTTGGAGGCCAAGGATGTGATCAC 1920
1921 CTGAGGTCAAGGATTTCAAAAACAGCTGSCCAACATGTGAAACCCCGTCTCTACTAAA 1980
1921 CTGAGGTCAAGGATTTCAAAAACAGCTGSCCAACATGTGAAACCCCGTCTCTACTAAA 1980
1981 ATACAAAAATTAGCTAGTGTGGTGGCGCATGCTGTAAATCCAGCTACTTTGGAGGGT 2040
1981 ATACAAAAATTAGCTAGTGTGGTGGCGCATGCTGTAAATCCAGCTACTTTGGAGGGT 2040
2041 GAGCAGAGAAATTTCTTAACCCCGGAGGAGGAGTGTGAGTGCAGTGAAGGAGATCACCA 2100
2041 GAGCAGAGAAATTTCTTAACCCCGGAGGAGGAGTGTGAGTGCAGTGAAGGAGATCACCA 2100
2101 CTGCACTCCAGCTGGGGGAGAGCGAGACTTCTCTCAAAAACCAAAAACAAAAAGA 2160
2101 CTGCACTCCAGCTGGGGGAGAGCGAGACTTCTCTCAAAAACCAAAAACAAAAAGA 2160
2161 ATTAAGCAAAATTAGCATTTGACAGAGAACCTGAAGGGGGTCAACCACTGACAGATTTTC 2220
2161 ATTAAGCAAAATTAGCATTTGACAGAGAACCTGAAGGGGGTCAACCACTGACAGATTTTC 2220
2221 TGTGCCACATGCCAAGTACTTTCTGAGGCATCACTGAGTGAAGTGTCCACATCTGAATCA 2280
2221 TGTGCCACATGCCAAGTACTTTCTGAGGCATCACTGAGTGAAGTGTCCACATCTGAATCA 2280
2281 TCCAGTCTTTGTTAGAACTTTTCAACCGGACAGGGAGCCAGGACTGGAATGCACTCTCT 2340
2281 TCCAGTCTTTGTTAGAACTTTTCAACCGGACAGGGAGCCAGGACTGGAATGCACTCTCT 2340
2341 GGTCACTGGCCAGAGAGTTGGCTTTGACCTTGAGACAGAGTGGCCAAACAAAGGAGCTGTT 2400
2341 GGTCACTGGCCAGAGAGTTGGCTTTGACCTTGAGACAGAGTGGCCAAACAAAGGAGCTGTT 2400
2401 AGTCTACTCTCCAGGAATCCAGGTGCTTCTCTCTGGAAGTGAATCATTTGGCCGAG 2460
2401 AGTCTACTCTCCAGGAATCCAGGTGCTTCTCTCTGGAAGTGAATCATTTGGCCGAG 2460
2461 CACTCCGATTTTTCTCTCTTCCAGGGAGAGATCTTAGGGCAGTATTTGGGAAAGACA 2520
2461 CACTCCGATTTTTCTCTCTTCCAGGGAGAGATCTTAGGGCAGTATTTGGGAAAGACA 2520
2521 TGGGCATGGAAGACACCGGGTGAATCATAGCTGCTGCTGCTGCTCATGTGTA 2580

2521 TGGGCATGGAAGACACCGGGTGAATCATAGCTGCTGCTGCTCATGTGTA 2580
2581 GGCTCTCTACAGACACCGAAAAAGATGGGGCAGAGGACAGATCAGTAGGGTCAGAGCATC 2640
2581 GGCTCTCTACAGACACCGAAAAAGATGGGGCAGAGGACAGATCAGTAGGGTCAGAGCATC 2640
2641 TCAGGACCGAGGGCAATATGCTTCTGAGCAGGGATTAAGAGCTTGGGCTCTCATATGTT 2700
2641 TCAGGACCGAGGGCAATATGCTTCTGAGCAGGGATTAAGAGCTTGGGCTCTCATATGTT 2700
2701 GTTTCCTGGGCTCAACTGCGCAGCTCCGTCACTTACTGTTGCTGTGACCATGGCAAGTTA 2760
2701 GTTTCCTGGGCTCAACTGCGCAGCTCCGTCACTTACTGTTGCTGTGACCATGGCAAGTTA 2760
2761 TTCCATCTCTCCATATCTCTTTCTCACTTTTAAATGGAATAATGGGGTACCACCTCC 2820
2761 TTCCATCTCTCCATATCTCTTTCTCACTTTTAAATGGAATAATGGGGTACCACCTCC 2820
2821 CAGGGTCACAGAGAGCTTTACAGAAAAAGATTTGAAATGGCTTGCAGTAATAATTC 2880
2821 CAGGGTCACAGAGAGCTTTACAGAAAAAGATTTGAAATGGCTTGCAGTAATAATTC 2880
2881 AATACTGCGCAGCTATCTTATTTCCACATCCAGCCCTTTCGCCCTGCTGCTGGGTGAAA 2940
2881 AATACTGCGCAGCTATCTTATTTCCACATCCAGCCCTTTCGCCCTGCTGCTGGGTGAAA 2940
2941 CACATGTCAGTGTCTTCTGACGGTTTCCAAAAAGATTTCCAAAAATTAACAACCTGCCAG 3000
2941 CACATGTCAGTGTCTTCTGACGGTTTCCAAAAAGATTTCCAAAAATTAACAACCTGCCAG 3000
3001 TCTGAAGATCTCCAAAAAGATTTCCAAAAAGATTTCCAAAAATTAACAACCTGCCAG 3060
3001 TCTGAAGATCTCCAAAAAGATTTCCAAAAAGATTTCCAAAAATTAACAACCTGCCAG 3060
3061 TGCCCGCCCGGGTCTGAAACAGGATGCGTGGCGCAGGACACACACACACAGCCCTG 3120
3061 TGCCCGCCCGGGTCTGAAACAGGATGCGTGGCGCAGGACACACACACACAGCCCTG 3120
3121 TGTGTGCGCCCGGAGTCCCGGTCCGGTCCAGACCGCGTGGCTGCTGGGGCGGGC 3180
3121 TGTGTGCGCCCGGAGTCCCGGTCCGGTCCAGACCGCGTGGCTGCTGGGGCGGGC 3180
3181 AGAGCATTGTTTCGAGCGCTACAGACCCCGCGCTGCGCCGGAGGGAGGGGGCT 3240
3181 AGAGCATTGTTTCGAGCGCTACAGACCCCGCGCTGCGCCGGAGGGAGGGGGCT 3240
3241 TCCCGCGTCCCAAGCTCCAGATCCTGGGGTGGCTGCCACGCTCCTGCCACGCGCTG 3300
3241 TCCCGCGTCCCAAGCTCCAGATCCTGGGGTGGCTGCCACGCTCCTGCCACGCGCTG 3300
3301 GGGGACCGGAAGACGGAGCGGAGATTTAGTGGTGGCGCCCGCCCGAGGGTTTCAACACT 3360
3301 GGGGACCGGAAGACGGAGCGGAGATTTAGTGGTGGCGCCCGCCCGAGGGTTTCAACACT 3360
3361 GTTTCCTGAGAACTTCCCGCAGTCCCAACCGCTTCTCCGTGTCGCCAGGGCGGGT 3420
3361 GTTTCCTGAGAACTTCCCGCAGTCCCAACCGCTTCTCCGTGTCGCCAGGGCGGGT 3420
3421 CTTGGGCTAGGCTCCCGCGCCCGAGCCCAACCGGGTCCCGAGCCCTTCCAGAGAGAAA 3480
3421 CTTGGGCTAGGCTCCCGCGCCCGAGCCCAACCGGGTCCCGAGCCCTTCCAGAGAGAAA 3480
3481 GTCCTCCGACCGGGATGCGCGGCGAGAGCCCGAGCGGGTGGAGAGAAAGCTGAGAGG 3540
3481 GTCCTCCGACCGGGATGCGCGGCGAGAGCCCGAGCGGGTGGAGAGAAAGCTGAGAGG 3540
3541 AGAAACAGAGGGGAGGGGAGCGAGGAGCTGGCGGAGAGGGAAACAGCATTTCCGCCGA 3600
3541 AGAAACAGAGGGGAGGGGAGCGAGGAGCTGGCGGAGAGGGAAACAGCATTTCCGCCGA 3600
3601 GCCAATGGCAACCGCAGGACGAGGTGSCAAATTCCTTCCGCCAATTCAGCAGCGGA 3660
3601 GCCAATGGCAACCGCAGGACGAGGTGSCAAATTCCTTCCGCCAATTCAGCAGCGGA 3660

```

Db      3601  GCCAATGCGCAACGGCAGGAGGTGGCACCACAAATTCCTTCGGCCAAATGACGACCGCGA 3660
Qy      3661  GTTTACAGAAGCCTCATTTAGCATTTCCCGCAGAGGCGAGGGCAGAGGCCCGGTGG 3720
Db      3661  GTTTACAGAAGCCTCATTTAGCATTTCCCGCAGAGGCGAGGGCAGAGGCCCGGTGG 3720
Qy      3721  TGTGTTGTCGGTGTGGCAGCATCCCGCGCCCTTGTGCTGCGGTGCGCGAGCCTCGGCC 3780
Db      3721  TGTGTTGTCGGTGTGGCAGCATCCCGCGCCCTTGTGCTGCGGTGCGCGAGCCTCGGCC 3780
Qy      3781  TCTGTCCTCTCCCTCCCGCCCTTACCTTCCACGCGGACCGCGCGCCAGTCAACTCC 3840
Db      3781  TCTGTCCTCTCCCTCCCGCCCTTACCTTCCACGCGGACCGCGCGCCAGTCAACTCC 3840
Qy      3841  TCGCACATTTGCCCTTGTGTCAGCGGATAAAGGGGCTCAGGAATACCGGACACGGT 3900
Db      3841  TCGCACATTTGCCCTTGTGTCAGCGGATAAAGGGGCTCAGGAATACCGGACACGGT 3900
Qy      3901  CACCGTTGCCAGCTCTAGCTTTAAATTCCTCGGCTCGGGACCTCCACGACCGCGG 3958
Db      3901  CACCGTTGCCAGCTCTAGCTTTAAATTCCTCGGCTCGGGACCTCCACGACCGCGG 3958

RESULT 3
AC018783/c
LOCUS
DEFINITION Homo sapiens clone RP11-2M13, WORKING DRAFT SEQUENCE, 12 unordered
            pieces.
ACCESSION AC018783
VERSION   AC018783.3 GI:7229991
KEYWORDS  HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 159351)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
           Roguslavskiy,L., Bouckgalter,B., Brown,A., Burkett,G., Castle,A.,
           Choepel,Y., Collangelo,M., Collins,S., Collamore,A., Cooke,P.,
           Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
           Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,J.,
           Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
           Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
           MacDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
           Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H.,
           O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K., Pierre,N.,
           Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,
           Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
           Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
           Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.
           and Zody,M.
JOURNAL   Direct Submission
COMMENT   Submitted (19-DEC-1999) Whitehead Institute/MIT Center for Genome
           Research, 320 Charles Street, Cambridge, MA 02141, USA
           On Mar 12, 2000 this sequence version replaced gi:6692332.
           All repeats were identified using RepeatMasker:
           Smit, A.F.A. & Green, P. (1996-1997)
           http://ftp.genome.washington.edu/RM/RepeatMasker.html
           ----- Genome Center
           Center: Whitehead Institute/ MIT Center for Genome Research
           Center code: WIBR
           Web site: http://www-seq.wi.mit.edu
           Contact: sequence_submissions@genome.wi.mit.edu
           ----- Project Information
           Center project name: L2667
           Center clone name: 2_M13
           ----- Summary Statistics

```

```

Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 151339 bases at least Q40
Consensus quality: 155177 bases at least Q30
Consensus quality: 157011 bases at least Q20
Insert size: 167000; agarose-fp
Insert size: 158251; sum-of-contigs
Quality coverage: 4.9 in Q20 bases; agarose-fp
Quality coverage: 5.2 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

* 1
* 2297: contig of 2297 bp in length
* 2298
* 2397: gap of 100 bp
* 2398
* 5719: contig of 3322 bp in length
* 5720
* 5819: gap of 100 bp
* 5820
* 10149: contig of 4330 bp in length
* 10150
* 10249: gap of 100 bp
* 10250
* 16496: contig of 6247 bp in length
* 16497
* 16596: gap of 100 bp
* 16597
* 21823: contig of 5227 bp in length
* 21824
* 21924: gap of 100 bp
* 21924
* 25747: contig of 3824 bp in length
* 25748
* 31971: contig of 6124 bp in length
* 31972
* 32071: gap of 100 bp
* 32072
* 44449: contig of 12378 bp in length
* 44450
* 44549: gap of 100 bp
* 44550
* 64978: contig of 20429 bp in length
* 64979
* 65079: gap of 100 bp
* 65079
* 91491: contig of 26413 bp in length
* 91492
* 91591: gap of 100 bp
* 91592
* 121937: contig of 30346 bp in length
* 121938
* 122037: gap of 100 bp
* 122038
* 159351: contig of 37314 bp in length.

```

FEATURES

source

```

1..159351
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-2M13"
/clone_lib="RP11-11 Human Male BAC"

```

misc_feature

```

1..2297
/note="assembly_fragment"
2398..5719
/note="assembly_fragment"

```

misc_feature

```

5820..10149
/note="assembly_fragment"

```

misc_feature

```

10250..16496
/note="assembly_fragment"

```

misc_feature

```

16597..21823
/note="assembly_fragment"

```

misc_feature

```

21924..25747
/note="assembly_fragment"

```

misc_feature

```

vector_side:right
25848..31971
/note="assembly_fragment"

```

misc_feature

```

vector_side:left
32072..44449
/note="assembly_fragment"

```

misc_feature

```

44550..64978
/note="assembly_fragment"

```

misc_feature

```

65079..91491
/note="assembly_fragment"

```

[illegible]

QY	2043	GGCAGGGAATTTCTTGAACCCGGGAGCGAGAGTTGCAGTGAAGCGAGATCACCACT	2102
DB	57350	GGCAGGGAATTTCTTGAACCCGGGAGCGAGAGTTGCAGTGAAGCGAGATCACCACT	57291
QY	2103	GCACCTCCAGCCTGGGGAGAGCGAGACTTCCTCTCAAAAAAACAAGAAAT	2162
DB	57290	GCACCTCCAGCCTGGGGAGAGCGAGACTTCCTCTCAAAAAAACAAGAAAT	57231
QY	2163	TAAGCAAAATTAGACAATTTGCAGAGAGAACCTTGAAGGGGCTCAGACCACGTACAGATTTCTG	2222
DB	57230	TAAGCAAAATTAGACAATTTGCAGAGAGAACCTTGAAGGGGCTCAGACCACGTACAGATTTCTG	57171
QY	2223	TGCCACATGCCAAGTAATTTCTGAGGCATGACTGGATGAGCTGTCCATCTGAAATCATC	2282
DB	57170	TGCCACATGCCAAGTAATTTCTGAGGCATGACTGGATGAGCTGTCCATCTGAAATCATC	57111
QY	2283	CAGTCTTTGTTTCAGAACTTTCACACCGGACAGGAGCCAGGACTGGAATGCACTCTCTGG	2342
DB	57110	CAGTCTTTGTTTCAGAACTTTCACACCGGACAGGAGCCAGGACTGGAATGCACTCTCTGG	57051
QY	2343	TCACGTGCCACAGAGATTGGCTTTGACCCCTCAGACCAGTGGCCAAACAAAGGAGCTGCTTAG	2402
DB	57050	TCACGTGCCACAGAGATTGGCC-TGACCCCTCAGACCAGTGGCCAAACAAAGGAGCTGCTTAG	56992
QY	2403	TCTACCTCCAGAAATCCCAAGTGCTGTCTCTCTGGGAAAGTAATCAATTTGGCGCAGCA	2462
DB	56991	TCTACCTCCAGAAATCCCAAGTGCTGTCTCTCTCTGGGAAAGTAATCAATTTGGCGCAGCA	56932
QY	2463	CTCGTATTTCTTCCTCCAGGAGGAGGATCTCTAGGGCAGTATTTGGGAAGAAGCATG	2522
DB	56931	CTCGTATTTCTTCCTCCAGGAGGAGGATCTCTAGGGCAGTATTTGGGAAGAAGCATG	56872
QY	2523	GGCATGGAAGGACACCGGGTGAATGATACCTGCTGCTGTCTGAGCTCTCATGTGAAGG	2582
DB	56871	GGCATGGAAGGACACCGGGTGAATGATACCTGCTGCTGTCTGAGCTCTCATGTGAAGG	56812
QY	2583	CTCCTACAGACACGGAAGATGGGGGACAGGACAGATCAGTAGGCTCAGAGCATCTC	2642
DB	56811	CTCCTACAGACACGGAAGATGGGGGACAGGACAGATCAGTAGGCTCAGAGCATCTC	56752
QY	2643	AGGACCGAGGGCAATATGCTCTGAGCAGGAGTAAGAGCTTGGGCTCATATGTGT	2702
DB	56751	AGGACCGAGGGCAATATGCTCTGAGCAGGAGTAAGAGCTTGGGCTCATATGTGT	56692
QY	2703	TTCTGGGCTCAACTGCCAGCTCCGTCACTACTGCTGTGTGACCATGGGCAAGTATT	2762
DB	56691	TTCTGGGCTCAACTGCCAGCTCCATCACTTACTGTGTGTGACCATGGGCAAGTATT	56632
QY	2763	CCATCTCTCGATATCTCTTCTCTCACTTTTAAATTGGAAATATGGGGTACCCACTCCCA	2822
DB	56631	CCATCTCTCGATATCTCTTCTCTCACTTTTAAATTGGAAATATGGGGTACCCACTCCCA	56572
QY	2823	GGGTACAGAGAGGCTTACAGAAACGATCTGTGGAATGGCTTGCAGTAATTAATTCAA	2882
DB	56571	GGGTACAGAGAGGCTTACAGAAACGATCTGTGGAATGGCTTGCAGTAATTAATTCAA	56512
QY	2883	TACCTGCCAGCTATTCTTATTCACATCCAAAGCCCTTTTCGGCTGCTGCTGGGTGAAACA	2942
DB	56511	TACCTGCCAGCTATTCTTATTCACATCCAAAGCCCTTTTCGGCTGCTGCTGGGTGAAACA	56452
QY	2943	CATGTCAAGTGTTCCTGACGGTTTCCAAAGAAGATTCCAAAATTAACAATTCAGCTC	3002
DB	56451	CATGTCAAGTGTTCCTGACGGTTTCCAAAGAAGATTCCAAAATTAACAATTCAGCTC	56392
QY	3003	TGAAGAAATCTCCAAACATCCCGACGATCTCTGAGGCGGGGTTGGGATGGGACTG	3062
DB	56391	TGAAGAAATCTCCAAACATCCCGACGATCTCTGAGGCGGGGTTGGGATGGGACTG	56332
QY	3063	CCGCCCCGGGTCTTGAACAGGATGGTGGCGGAGGCACACACACAGCCAGCTGTG	3122
DB	56331	CCGCCCCGGGTCTTGAACAGGATGGTGGCGGAGGCACACACACAGCCAGCTGTG	56272
QY	3123	TGTGGGCGGAGTCCGGTGGGTCGGGTCGGGTTGACAGCGCGTGGCTGGTGGGCGGGCAG	3182

Db	56271																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
----	-------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

COMMENT

requests: clonerequest@sanger.ac.uk
 On Sep 19, 2001 this sequence version replaced gi:14586033.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
 database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 10, constructed by the Sanger Centre Chromosome 10
 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr10>
 RP11-34D15 is from the library RP11-11.1 constructed by the group
 of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACe3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-34D15. It may be shorter because we sequence overlapping
 sections only once, except for a short overlap.
 The true right end of clone RP11-34D15 is at 115278 in this
 sequence. The true left end of clone RP11-285F16 is at 56886 in
 this sequence. The true right end of clone RP11-316M21 is at 2000
 in this sequence.

FEATURES

Location/Qualifiers
 1..115278
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP11-34D15"
 /clone_lib="RP11-11.1"

ORIGIN

Query Match 99.0%; Score 3917.8; DB 9; Length 115278;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3952; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
 3 GAATTCATCCATTTAAATCATACATATTTAATGGCTTTTAGTATATATCATAGTTGTGCAT 62
 64592 GAATTCATCCATTTAAATCATACATATTTAATGGCTTTTAGTATATATCATAGTTGTGCAT 64651
 63 CCATCACATCCATTTAGAACAGTTTATATCTCCAAATAATAAACCTCGATTCCTTAG 122
 64652 CCATCACATCCATTTAGAACAGTTTATATCTCCAAATAATAAACCTCGATTCCTTAG 64711
 123 CCATCACATCCATTCCTCCTCTCCAAAGCCCTGGGCAACCACTACTT 182
 64712 CCATCACATCCATTCCTCCTCTCCAAAGCCCTGGGCAACCACTACTT 64771
 183 TCTGCTCTAATAATTTGCCAATCTGGACATTTATATTAATGAAGCAACCAATG 242
 64772 TCTGCTCTAATAATTTGCCAATCTGGACATTTATATTAATGAAGCAACCAATG 64831
 243 GAGACTTTGTGACTGGCTGCTTTCACTTAGCATTTCTATTTTAAAGGCTCATTTAGTTACA 302
 64832 GAGACTTTGTGACTGGCTGCTTTCACTTAGCATTTCTATTTTAAAGGCTCATTTAGTTACA 64891
 303 GTACTTAGCAGTACTTCACTCTTTTATTTTATCTCAATGGTATTCACCTGTGTGGGTATCC 362
 64892 GTACTTAGCAGTACTTCACTCTTTTATTTTATCTCAATGGTATTCACCTGTGTGGGTATCC 64951

QY	363	CATATCATATTTATTAGACAGGTTCTCACTCTGTACCCAGGCTGGAGTGCATGGCAC	422
Db	64952	CATATCATATTTATTAGACAGGTTCTCACTCTGTACCCAGGCTGGAGTGCATGGCAC	65011
QY	423	AATCATAGCTCACTGTAAACCTCAAACTCTGGGCTCAAGTGATCTACTACTCAGGCTC	482
Db	65012	AATCATAGCTCACTGTAAACCTCAAACTCTGGGCTCAAGTGATCTACTACTCAGGCTC	65071
QY	483	CAGAGTAGCTAGACATACAGGCAACACAGCCATACCTGGCTAATTTTTTTTAAATTT	542
Db	65072	CAGAGTAGCTAGACATACAGGCAACACAGCCATACCTGGCTAATTTTTTTTAAATTT	65131
QY	543	TCAATTTATGTATTCATTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	602
Db	65132	TCAATTTATGTATTCATTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	65191
QY	603	TTTGTACCAGGCTGGAGGCGAGTGGCATGTGGTGACAGCTGAGCAGCTTGACTTCTCTGG	662
Db	65192	TTTGTACCAGGCTGGAGGCGAGTGGCATGTGGTGACAGCTGAGCAGCTTGACTTCTCTGG	65251
QY	663	GCTCAAGTATCTCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGTGTCAAC	722
Db	65252	GCTCAAGTATCTCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGTGTCAAC	65311
QY	723	ATGCGCTGGCTGATATTTTTTTTCTTTGAAACAGGATATCACTCTGTTGCCAGGCTGGAGT	782
Db	65312	ATGCGCTGGCTGATATTTTTTTTCTTTGAAACAGGATATCACTCTGTTGCCAGGCTGGAGT	65371
QY	783	ACAGTGGGCTAATAATAGCTCACTGAGCCTCCCTCTGGGCTCAAGCAATCCGCTGGC	842
Db	65372	ACAGTGGGCTAATAATAGCTCACTGAGCCTCCCTCTGGGCTCAAGCAATCCGCTGGC	65431
QY	843	CTCAGATCTCTGAGTAGCTGGGACTACAGGCTTTGTGCCACAGGCCCGAGTAAAGTTTAA	902
Db	65432	CTCAGATCTCTGAGTAGCTGGGACTACAGGCTTTGTGCCACAGGCCCGAGTAAAGTTTAA	65491
QY	903	AAATGATTTTGGTATAGAGGAGTCTTGCTATCTGCTGAGGCTGTATTTTATCTGT	962
Db	65492	AAATGATTTTGGTATAGAGGAGTCTTGCTATCTGCTGAGGCTGTATTTTATCTGT	65551
QY	963	GAGACAGGCTCTCACTATGTTGCCATGATCCGCCATCCCACTTCCCAAGTGTCTATC	1022
Db	65552	GAGACAGGCTCTCACTATGTTGCCATGATCCGCCATCCCACTTCCCAAGTGTCTATC	65611
QY	1023	TTATCTGTTCAATAGTACAGATTTAGGCTTTGTTTCCACTTTTACCAATATG	1082
Db	65612	TTATCTGTTCAATAGTACAGATTTAGGCTTTGTTTCCACTTTTACCAATATG	65671
QY	1083	AATAATCTCCAGTGAATATTCATGTATACATTTTGTGTGGCATATGTTTTCATTTCTGT	1142
Db	65672	AATAATCTCCAGTGAATATTCATGTATACATTTTGTGTGGCATATGTTTTCATTTCTGT	65731
QY	1143	TGGGTTTATATCTAGGAGTGAATTTGCTGGATCCCGGTTAATATTTTACAGGAGGTT	1202
Db	65732	TGGGTTTATATCTAGGAGTGAATTTGCTGGATCCCGGTTAATATTTTACAGGAGGTT	65791
QY	1203	CAGGGGAAGAAAACCTTGGGAAATGAAGCATGTTTATAGAAATCAGCAAGTGCAGGGGT	1262
Db	65792	CAGGGGAAGAAAACCTTGGGAAATGAAGCATGTTTATAGAAATCAGCAAGTGCAGGGGT	65851
QY	1263	TTTTCGAGTTTATTTTATTTTCTGTTGACAAATGTGCAATTTGATGAAGATCAAGTT	1322
Db	65852	TTTTCGAGTTTATTTTATTTTCTGTTGACAAATGTGCAATTTGATGAAGATCAAGTT	65911
QY	1323	ATACTAAGTGAAGTGAAGTAAAGCTGAATAGGCGTTTACAGTAAATCATGAAG	1382
Db	65912	ATACTAAGTGAAGTGAAGTAAAGCTGAATAGGCGTTTACAGTAAATCATGAAG	65971
QY	1383	CACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATAAATAACCAATT	1442
Db	65972	CACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATAAATAACCAATT	66031
QY	1443	TTTTTTTTTTTTTTTGAGAAAGAGTCTTGCTCTTTTCACTCTGGCTGGAGGGGAGTGGTGTG	1502

Db	66032	TTTTTTTTTTTTTTTTCAGAAAGAGTCTTGCTCTTTTACCCCTGGCTGGAGGCGAGTGGTGTG	66091
QY	1503	ATCTCAGCTCACTGCAAACTTTTCGCTCCCGGGTTCAGAGCAATTCCTCGTTCAGCCCTCC	1562
Db	66092	ATCTCAGCTCACTGCAAACTTTTCGCTCCCGGGTTCAGAGCAATTCCTCGTTCAGCCCTCC	66151
QY	1563	CAAGTAGCTGGGATACAGGCACTTTCCACCAATGCCCAGCTGATTTTGTATTTTAGTA	1622
Db	66152	CAAGTAGCTGGGATACAGGCACTTTCCACCAATGCCCAGCTGATTTTGTATTTTAGTA	66211
QY	1623	GAGATGGGATTCACCTTTTGGCCCAAGCTGGTCTCAAACTTTTGGCTGTGCTGCTGCTG	1682
Db	66212	GAGATGGGATTCACCTTTTGGCCCAAGCTGGTCTCAAACTTTTGGCTGTGCTGCTGCTG	66271
QY	1683	GTAACATATTTGCTGCTGAGGTAGGGCCCCAGACCCCAAAATAAATAATCTTAGAA	1742
Db	66272	GTAACATATTTGCTGCTGAGGTAGGGCCCCAGACCCCAAAATAAATAATCTTAGAA	66331
QY	1743	TCCAAATCAGTGTGTTGTTGACCACTGTCACTTGAGAACCAAGTGTGACAGGGGCT	1802
Db	66332	TCCAAATCAGTGTGTTGTTGACCACTGTCACTTGAGAACCAAGTGTGACAGGGGCT	66391
QY	1803	CAGAGTAGAGTGTCTCTGCTCGAAGAGAAATAGAAATAATTTCCCGGGCCAG	1862
Db	66392	CAGAGTAGAGTGTCTCTGCTCGAAGAGAAATAGAAATAATTTCCCGGGCCAG	66451
QY	1863	GCCTGGTGGCTCATGCCCTGTAATCCAGCACTTTGGGAGGCCAAGGCATGTGGATCACT	1922
Db	66452	GCCTGGTGGCTCATGCCCTGTAATCCAGCACTTTGGGAGGCCAAGGCATGTGGATCACT	66511
QY	1923	GAGTCAAGAGTTCAAAACACGCTGCGCAACATFGGTGAAACCCCGTCTCTACTAAAAAT	1982
Db	66512	GAGTCAAGAGTTCAAAACACGCTGCGCAACATFGGTGAAACCCCGTCTCTACTAAAAAT	66571
QY	1983	ACAAAAATTAGCTAAGTGTGGCGCTGCTGTAATCCAGCTACTTTGGGAGGTGA	2042
Db	66572	ACAAAAATTAGCTAAGTGTGGCGCTGCTGTAATCCAGCTACTTTGGGAGGTGA	66631
QY	2043	GGCAGGAGAAATTTCTTGAACCCGGAGGCGAGAGTTCAGTGAAGCGAGATCACCACT	2102
Db	66632	GGCAGGAGAAATTTCTTGAACCCGGAGGCGAGAGTTCAGTGAAGCGAGATCACCACT	66691
QY	2103	GCACTCAGCTGGGGAGAGAGCGAGATTCCTCTCAAAAAACAAAAACAAAAAGAAAT	2162
Db	66692	GCACTCAGCTGGGGAGAGAGCGAGATTCCTCTCAAAAAACAAAAACAAAAAGAAAT	66751
QY	2163	TAAGCAAAATTAGACATTTGACAGAGBACCTGAGGGGGTTCAGACCAAGTACATTTCTG	2222
Db	66752	TAAGCAAAATTAGACATTTGACAGAGBACCTGAGGGGGTTCAGACCAAGTACATTTCTG	66811
QY	2223	TGCCACATGCCAAAGTACTTCTGAGGATGACTGGATGAGCTGTCACATCTGAAATCATC	2282
Db	66812	TGCCACATGCCAAAGTACTTCTGAGGATGACTGGATGAGCTGTCACATCTGAAATCATC	66871
QY	2283	CAGTCTTTTTCAGAACTTTTCACACCGGAGCGAGACTGGAAATGCACTCTCTCTG	2342
Db	66872	CAGTCTTTTTCAGAACTTTTCACACCGGAGCGAGACTGGAAATGCACTCTCTCTG	66931
QY	2343	TCACTGGCCAGAGTTGGCTTGACCTGAGACAGTGGCCCAACAAAGAGCTGTTAG	2402
Db	66932	TCACTGGCCAGAGTTGGCTTGACCTGAGACAGTGGCCCAACAAAGAGCTGTTAG	66990
QY	2403	TCTACCTCCAGGAAATCCAGGTGCTTGTCTTCTGGGAAGTGAATCATTTGGCGCAGCA	2462
Db	66991	TCTACCTCCAGGAAATCCAGGTGCTTGTCTTCTGGGAAGTGAATCATTTGGCGCAGCA	67050
QY	2463	CTCCGTATTTCTCTCTTCCAGGGGAAGGATCTTAGGGGAGTATTTGGGAAAGACATG	2522
Db	67051	CTCCGTATTTCTCTCTTCCAGGGGAAGGATCTTAGGGGAGTATTTGGGAAAGACATG	67110
QY	2523	GGCATGGGAAGGACACCGGGTGAATGATAGCTGCTGGTCTGAGCTCTCATGTGAGG	2582

Db	67111	GGCATGGGAAGGACACCGGGTGAATGATAGCTGCTGCTGCTGCTCTCATNGGTAAGG	67170
QY	2583	CTCCTACAGACACGGAAAAAGATGGGGGCAACAGGACAGATCAGTAGGGTCAGAGCATCTC	2642
Db	67171	CTCCTACAGACACGGAAAAAGATGGGGGCAACAGGACAGATCAGTAGGGTCAGAGCATCTC	67230
QY	2643	AGGACCGAGGGCAATATGGTCCCTGAGCAGGAGGATTAAGAGCTTGGGCTCTCATATGGTGT	2702
Db	67231	AGGACCGAGGGCAATATGGTCCCTGAGCAGGAGGATTAAGAGCTTGGGCTCTCATATGGTGT	67290
QY	2703	TTCTGGGCTCAACTGCCAGCTCCGTCACCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2762
Db	67291	TTCTGGGCTCAACTGCCAGCTCCATCACTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	67350
QY	2763	CCATCTCTCATATCT	2822
Db	67351	CCATCTCTCATATCT	67410
QY	2823	GGGTACACAGAGAGGCTTACAGAAAAAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2882
Db	67411	GGGTACACAGAGAGGCTTACAGAAAAAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	67470
QY	2883	TACCTGCCAGTATTTCTTTATTTCCACATCCCAAGCCCTTTCCGCTGCTGCTGCTGCTGCTGCT	2942
Db	67471	TACCTGCCAGTATTTCTTTATTTCCACATCCCAAGCCCTTTCCGCTGCTGCTGCTGCTGCTGCT	67530
QY	2943	CATGTCACTGTTTCTGACGGTTTCCAAAGAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3002
Db	67531	CATGTCACTGTTTCTGACGGTTTCCAAAGAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT	67590
QY	3003	TGAAGAAATCTCCAAAAACATCCCGCAGCATCTCTGAGGGCGCGGCTTGGGGATGGGACTG	3062
Db	67591	TGAAGAAATCTCCAAAAACATCCCGCAGCATCTCTGAGGGCGCGGCTTGGGGATGGGACTG	67650
QY	3063	CCCGCCGGGTCTCTGAACAGGATGCGTGGCGGCGAGGACACACACACACACACACACACACAC	3122
Db	67651	CCCGCCGGGTCTCTGAACAGGATGCGTGGCGGCGAGGACACACACACACACACACACACACAC	67710
QY	3123	TGTTCCGCGCGGAGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGG	3182
Db	67711	TGTTCCGCGCGGAGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGG	67770
QY	3183	AGCCATTTGTCGAGAGCGTACCGAGCCCCCGCGCTCGCCCGGAGGAGGCGGGGCTTC	3242
Db	67771	AGCCATTTGTCGAGAGCGTACCGAGTCCCGCGCTCGCCCGGAGGAGGCGGGGCTTC	67830
QY	3243	CCGCGT - CCCCAGTCTCAGATCTCTGGGGTGGTGCACGCTCTCCCTGCGCACCGCGCTGG	3301
Db	67831	CCGCGTCCCGCAAGCTCCAGATCTCTGGGGTGGTGCACGCTCTCCCTGCGCACCGCGCTGG	67890
QY	3302	GGGGACGGGAAGACGGGACGGGAGATGTTAGTGGTGGGCGCCCCCGAGGGTTCCACACTG	3361
Db	67891	GGGGAC - GGAAACGGGACGGGAGATGTTAGTGGTGGGCGCCCCCGAGGGTTCCACACTG	67949
QY	3362	TTTTCTTGAGAAATTTCCCGAGTGGCCACCCACCCGCTTCCTCGGTGTCGCCGAGGGCGGCTC	3421
Db	67950	TTTTCTTGAGAAATTTCCCGAGTGGCCACCCACCCGCTTCCTCGGTGTCGCCGAGGGCGGCTC	68009
QY	3422	CTGGGCTAGGCTCCGGGCCCCAGCCCCAAAACCGGGTCCCGAGCCCCCTTCCAGAGAGAAAG	3481
Db	68010	CTGGGCTAGGCTCCGGGCCCCAGCCCCAAAACCGGGTCCCGAGCCCCCTTCCAGAGAGAAAG	68069
QY	3482	CTCCCGACCGGGGATCCCGGACAGGCGCCAGCGGGGTGGAAGAGAGCTGAGAGGA	3541
Db	68070	CTCCCGACCGGGGATCCCGGACAGGCGCCAGCGGGGTGGAAGAGAGCTGAGAGGA	68129
QY	3542	GAAACAGAGGGGAGGGGAGCGAGAGCTGGCGGACAGAGGGAACAGCAGATTCGCGCCGAG	3601
Db	68130	GAAACAGAGGGGAGGGGAGCGAGAGCTGGCGGACAGAGGGAACAGCAGATTCGCGCCGAG	68189
QY	3602	CCATGGGCAACGGCAGGAGCAGAGGTGGCAACCAATTTCCCTTCCGCGCAATGACGAGCGGAG	3661
Db	68190	CCATGGGCAACGGCAGGAGCAGAGGTGGCAACCAATTTCCCTTCCGCGCAATGACGAGCGGAG	68249

3662 TTTACAGAAGCTCATTAGCAATTTCCACAGAGGAGGGGACAGGGGACAGAGCCGGGTGGT 3721
68250 TTTACAGAAGCTCATTAGCAATTTCCACAGAGGAGGGGACAGGGGACAGAGCCGGGTGGT 68309
3722 GTGTGTGCGGTGCGGAGCATCCCGCGGCCCTGCTCGGTGCGCGGAGCCTCGGCGCT 3781
68310 GTGTGTGCGGTGCGGAGCATCCCGCGGCCCTGCTCGGTGCGCGGAGCCTCGGCGCT 68369
3782 CTGTCTCTCCCTCCCGCCCTTACTCCACGCGGACCGCCCGCCGCACTCACTCCT 3841
68370 CTGTCTCTCCCTCCCGCCCTTACTCCACGCGGACCGCCCGCCGCACTCACTCCT 68429
3842 CCGACTTGGCCCTGCTTGGCAGCGGATAAAGGGGCTGAGGAATACCGACACGCTC 3901
68430 CCGACTTGGCCCTGCTTGGCAGCGGATAAAGGGGCTGAGGAATACCGACACGCTC 68489
3902 ACCGTTGCGAGCTCAGGCTTTAAATCCCGCTCGGCGACTCCACGACCGCGG 3958
68490 ACCGTTGCGAGCTCAGGCTTTAAATCCCGCTCGGCGACTCCACGACCGCGG 68546

RESULT 5
AL391723/c
LOCUS
DEFINITION Homo sapiens chromosome 10 clone RP11-38B21, WORKING DRAFT
SEQUENCE, 14 unordered pieces.
ACCESSION AL391723
VERSION AL391723.7 GI:10186824
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Burton, J.
Direct Submission
Sanger Centre, Hinxton, Cambridgeshire,
Submitted (10-SEP-2001) humquery@sanger.ac.uk
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Sep 19, 2000 this sequence version replaced gi:10178830.

Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

Project Information
Center project name: BA38B21

Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 175002 bases at least Q40
Consensus quality: 179431 bases at least Q30
Consensus quality: 182156 bases at least Q20
Insert size: 184344; sum-of-contigs
Insert size: 140788; 6.0% error; agarose-fp
Quality coverage: 4.15x in Q20 bases; sum-of-contigs Quality
coverage: 6.51x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 59110: contig of 59110 bp in length
* 59111: gap of 100 bp
* 59211: contig of 2075 bp in length
* 61286: gap of 100 bp
* 61386: contig of 2994 bp in length

NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 59110: contig of 59110 bp in length
* 59111: gap of 100 bp
* 59211: contig of 2075 bp in length
* 61286: gap of 100 bp
* 61386: contig of 2994 bp in length

FEATURES

source
1..185644
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone_lib="RP11-38B21"
/clone_id="RP11-38B21.1"
1..59110
/note="assembly_fragment:02307"
clone_end:17
vector_side:left
59211..61285
/note="assembly_fragment:00197"
fragment_chain:1
61386..64379
/note="assembly_fragment:01916"
fragment_chain:1
64480..66747
/note="assembly_fragment:00079"
fragment_chain:1
66848..114355
/note="assembly_fragment:00585"
fragment_chain:2
114456..125406
/note="assembly_fragment:02383"
fragment_chain:2
125507..127673
/note="assembly_fragment:00150"
127774..130111
/note="assembly_fragment:00464"
130212..132354
/note="assembly_fragment:00629"
132455..136537
/note="assembly_fragment:00814"
136638..139127
/note="assembly_fragment:00861"
139228..181355
/note="assembly_fragment:02212"
181456..183524
/note="assembly_fragment:02327"
183625..185644
/note="assembly_fragment:02419"

ORIGIN

Query Match 46.2%; Score 1827.2; DB 2; Length 185644;
Best Local Similarity 93.9%; Pred. No. 0; Mismatches 118; Indels 3; Gaps 3;
Matches 1871; Conservative
1968 GTCCTACTAAATAACAAAAAATTAGTGTGTGGCGCGCTGTTATCCGACG 2027

[illegible]

DEFINITION	Papio hamadryas clone RP41-194B17, WORKING DRAFT SEQUENCE.		
ACCESSION	AC139668	Db	139426
VERSION	AC139668.1 GI:28273351		
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT.	Qy	2038
SOURCE	Papio hamadryas (hamadryas baboon)	Db	139486
ORGANISM	Papio hamadryas	Qy	2098
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Db	139546
	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;	Qy	2098
	Cercopitheciinae; Papio.	Db	139546
REFERENCE	1 (bases 1 to 161827)	Qy	2098
AUTHORS	Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,	Qy	2098
	Peng, Z., Malinov, I. and Rubin, E.M.	Db	139605
TITLE	Direct Submission	Qy	2158
JOURNAL	Unpublished	Db	139605
REFERENCE	2 (bases 1 to 161827)	Qy	2158
AUTHORS	Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,	Qy	2158
	Peng, Z., Malinov, I. and Rubin, E.M.	Db	139605
TITLE	Direct Submission	Qy	2158
JOURNAL	Submitted (08-FEB-2003) Genome Sciences, Lawrence Berkeley National	Qy	2158
	Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA	Db	139605
COMMENT	Sequence Produced by Berkeley PGA	Qy	2158
	Web site: http://pga.lbl.gov	Qy	2158
	Center Code: PGABERK	Qy	2158
	Center Project Name: B051	Qy	2158
	Bac Clone Name: RP41-194B17	Qy	2158
	This sequence has been compared to sequences of other species	Qy	2158
	using Vista (http://www.gsdl.gov/vista/). The results can be	Qy	2158
	viewed at:	Qy	2158
	http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=SCD	Qy	2158
	The order-orientation of the draft sequence was accomplished by	Qy	2158
	using:	Qy	2158
	Avid (http://baboon.math.berkeley.edu/mavid/),	Qy	2158
	Lagan (http://lagan.stanford.edu/) and paired end information.	Qy	2158
	Funding agent: Programs for Genomic Applications (NHLBI)	Qy	2158
	Summary Statistics:	Qy	2158
	Sequencing vector: Plasmid; pUC18	Qy	2158
	Chemistry: Dye-terminator Big Dye	Qy	2158
	Assembly program: Phrap version 0.990329.	Qy	2158
	* NOTE: This is a 'working draft' sequence. It currently	Qy	2158
	* consists of 1 contigs. The true order of the pieces	Qy	2158
	* is not known and their order in this sequence record is	Qy	2158
	* arbitrary. Gaps between the contigs are represented as	Qy	2158
	* runs of N, but the exact sizes of the gaps are unknown.	Qy	2158
	* This record will be updated with the finished sequence	Qy	2158
	* as soon as it is available and the accession number will	Qy	2158
	* be preserved.	Qy	2158
FEATURES	1 161827: contig of 161827 bp in length.	Qy	2158
source	Location/Qualifiers	Qy	2158
	1. 161827	Qy	2158
	/organism="Papio hamadryas"	Qy	2158
	/mol_type="genomic DNA"	Qy	2158
	/db_xref="taxon:9557"	Qy	2158
	/clone="RP41-194B17"	Qy	2158
ORIGIN		Qy	2158
	Query Match 44.4%; Score 1758; DB 2; Length 161827;	Qy	2158
	Best Local Similarity 92.4%; Pred No. 0;	Qy	2158
	Matches 1943; Conservative 0; Mismatches 135; Indels 24; Gaps 8;	Qy	2158
Qy	1858 GCCAGCGGTGGTCTGATGCTGTAATCCAGACATTTGGAGGCGCAGGATGGAT 1917	Qy	2158
Db	139308 GCTGGGACGGTGGTCTGATGCTGTAATCCAGACATTTGGAGGCGCAGGAT 139367	Qy	2158
Qy	1918 CACCTGAGTCTAGGAGTTCAAAACCGAGCTGCCACATGGTGAACCCGCTCTACTA 1977	Qy	2158
Db	139368 CA--TGAGGTCTAGGATGCAAGACCGAGCTGCCACATGGTGAACCCGCTCTACTA 139425	Qy	2158
Qy	1978 AATAATACAAAAAATTAGCTAAGTGTGGTGGCGCATGCTGTAATCCAGCTACTTTGGGAG 2037	Qy	2158

Db	140490	GACTGTTTGGCCCGGGGCTGAACAGGATTGCTGCGCTAGG--CACACACACACGACCCAGA	140547
QY	3118	CTGTGTGTGCGGCCGCGAGTCCGCTGCGGTCCGGGTGACAGCGCTGCTGTGGTGGGCGG	3177
Db	140548	C--TGCCTGCGGCCGAGTCCGCTGCGGTCCGGGTGACAGCGCTGCTGTGGTGGGCGG	140605
QY	3178	GGCAGAGCCATTGTTTCGCAGGCGTACCGAGCCCCCGCGCTCGCCCGGAGGAGGCGCGG	3237
Db	140606	GGCAGAGCCATTGTTTCGCAGGCGTACCTGAGCGTCTGAGGTGCGTCCCGGAGGAGGCGCGG	140665
QY	3238	GCCTCCCGCGTCCCG--AAGCTCCAGATCCTGGGTGGTGGCACTGTCTCCCTGCGCACGCG	3296
Db	140666	GCCTCCCGCGTCCCGGAGCTCCCGATCCCGGGTGGCGGCTGCTCTCCCTGCGCACGCG	140725
QY	3297	CCTGGGGGACGGGNAGACGGGACGGAGATGTTAGTGTGTGGGCGCCCCCGAGGGTTCAC	3356
Db	140726	CCTGGGGGACGGGNAAGACGGGACGGGAATTGTGTAGTGTGGCGTCTCCGAGGCTTCCC	140785
QY	3357	CACCTGTTTCTTGAGAAACTTCCCGAGTGCCACCCACCCGTTCTCGTGTGCCCGAGGCG	3416
Db	140786	CACCTGTTTCTTGAGAAACTTCCCGAGTGCCACCCACCCGTTCTCGCGCGTCCCGGGG	140845
QY	3417	CGGCTCTGGGCTAGGCTCCGGCCCCAGCCCCCAACCCGGGTCCTCCAGCGCCCTTCAGAGA	3476
Db	140846	CGGCTCTGGGCTAGACTCCGGCCCCAGCCCCACACCCGGGTCCTCCAGTCCCTTCAGAGA	140905
QY	3477	GAAAGCTCCGACGCGGATGCCGGGACAGAGCCGAGCGGGGTGGGAAGAACTGAG	3536
Db	140906	GAAAGCTCCTTAACGCGGGGTGCCGGGACAGAGCCGAGCGGGGTGGGAAGAACTGAG	140965
QY	3537	AAGAGAAACAGAGGGGAGGGGACGAGGAGCTGGCGGACAGAGGAAACAGAGATTGCG	3596
Db	140966	AAGAGAAAGGAGGGGAGGGGAGCGAGAGCTGGCGGACAGAGGAAACAGAGATTGCG	141025
QY	3597	CGAGCCCAATGCAACGGCAGAGCAGGTGTGCACCAATTCCTTCGGCCATGACGAC	3656
Db	141026	CGAGCCCAATGCAACTGCAGGACGAGGTGGCAACCAATTCCTTCGGCCATGACGAC	141085
QY	3657	CGAGTTTACAGAGCCTATTAGCATTTCCCGAGGACGAGGCGAGGCGGAGAGCCGG	3716
Db	141086	CGAGTTTACAGAGCCTATTAGCATTTCCCGAGGACGAGGCGAGGCGGAGGCGGCGG	141145
QY	3717	GTGTGTGTGTGTCGGTGTGCGGAGCATCCCGCGCCCTGCTGCGGTCCCGAGCGCTC	3776
Db	141146	GTGTGTGTGTGTCGGTGTGCGGAGCAT--CCGCGCCCTGCTGCGGTCTGCGAGCCTT	141204
QY	3777	GGCCTGTGTCTCCCTCCCGCCCTTACCTCCAGCGGGNACCGCCCGCCAGTCAAA	3836
Db	141205	GGCCTGTGTCTCCCTCCCGCCCTTACCTCCAGCGGGNACCGCCCGCCAGTCAAA	141264
QY	3837	CTCCTCGCACTTTGCCCTGCTTGGCAGCGGATAAAAGGGGGCTGAGGAAATACCGACA	3896
Db	141265	CTCCTCGCACTTTGCCCTGCTTGGCAGCGGATAAAAGGGAGCTGAGGAAATACCGACA	141324
QY	3897	CGGTACCCGTTTGCACAGCTTACCTTTTAAATTCGCGGTTCGGGAGCTTCACGACCCG	3956
Db	141325	CGGTACCCCAATGCCAGCTCTAGCCCTTAAATTCGCGGTTCGGGAGCTTCACGACCCG	141384
QY	3957	GG 3958	
Db	141385	GG 141386	

RESULT 8					
AC026883					
LOCUS					
DEFINITION	AC026883	179879 bp	DNA	linear	HTG 29-MAY-2000
	Homc sapiens chromosome 3 clone RP11-799L19 map 3p,				WORKING DRAFT
	SEQUENCE, 41 unordered pieces.				
ACCESSION	AC026883				
VERSION	AC026883.3	GI:8101285			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.				
SOURCE	Homc sapiens (human)				
ORGANISM	Homc sapiens				

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominoide; Homo.

1 (bases 1 to 179879)

Bao, J., Bao, Q., Rao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Guan, Q., Gu, X., Guo, D., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, G., Li, J., Li, L., Li, Y., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, S., Sun, M., Sun, W., Sun, Y., Tao, R., Wang, H., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Yu, J. and Yang, H.

Chromosome 3p genomic sequence

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

2 (bases 1 to 179879)
Li, W., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W.,
Sun, X., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y., Qi, X.,
Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y., Li, G.,
Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D.,
Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J.
and Yang, H.

TITLE

JOURNAL

Direct submission
Submitted (25-MAR-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
On MAR-20, 2000 this sequence replaced g1:7417472.

COMMENT

```

-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hg.igtp.ac.cn
http://www.genomics.org.cn
Contact:hg@igtp.ac.cn
----- Project Information
Center project name:rlk project
Center clone name: Rpl1-799L19
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator ET 55% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 149369 bases at least Q40
Consensus quality: 160276 bases at least Q30
Consensus quality: 167926 bases at least Q20
Insert size: 148122; sum-of-contigs
Quality coverage: 3.56x in Q20 bases;sum-of-contigs

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 41 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	*	1	1496:	contig of 1496 bp in length
*	*	1497	1596:	gap of unknown length
*	*	1597	2949:	contig of 1353 bp in length
*	*	2950	3409:	gap of unknown length
*	*	3050	4409:	contig of 1360 bp in length
*	*	4410	4509:	gap of unknown length
*	*	4510	6137:	contig of 1628 bp in length
*	*	6138	6237:	gap of unknown length
*	*	6238	7384:	contig of 1047 bp in length
*	*	7285	7384:	gap of unknown length
*	*	7385	9013:	contig of 1629 bp in length
*	*	9014	9113:	gap of unknown length
*	*	9114	10444:	contig of 1331 bp in length
*	*	10445	10544:	gap of unknown length
*	*	10545	12153:	contig of 1609 bp in length
*	*	12154	12253:	gap of unknown length
*	*	12254	14386:	contig of 2133 bp in length
*	*	14387	14386:	gap of unknown length

14487 16990: contig of 2504 bp in length
14991 17090: gap of unknown length
17091 19022: contig of 1932 bp in length
19023 19122: gap of unknown length
19123 20497: contig of 1375 bp in length
20498 20597: gap of unknown length
20598 22009: contig of 1612 bp in length
22010 22309: gap of unknown length
22310 24677: contig of 2368 bp in length
24678 24777: gap of unknown length
24778 27655: contig of 2878 bp in length
27656 30189: contig of 2434 bp in length
30190 30289: gap of unknown length
30290 33658: contig of 3369 bp in length
33659 33758: gap of unknown length
33759 36912: contig of 3154 bp in length
36913 37012: gap of unknown length
37014 41069: contig of 4057 bp in length
41070 41169: gap of unknown length
41170 43658: contig of 2489 bp in length
43659 43758: gap of unknown length
43759 46574: contig of 2816 bp in length
46575 46674: gap of unknown length
46676 49890: contig of 3216 bp in length
49891 53751: contig of 3761 bp in length
53752 53851: gap of unknown length
53852 57378: contig of 3527 bp in length
57379 57478: gap of unknown length
57479 61598: contig of 4120 bp in length
61599 61698: gap of unknown length
61699 65686: contig of 3988 bp in length
65687 65786: gap of unknown length
65787 70673: contig of 4887 bp in length
70674 70773: gap of unknown length
70774 76442: contig of 5669 bp in length
76443 80193: contig of 3651 bp in length
80194 80293: gap of unknown length
80294 83630: contig of 3337 bp in length
83631 83730: gap of unknown length
83731 89196: contig of 5466 bp in length
89197 89296: gap of unknown length
89297 93491: contig of 4195 bp in length
93492 93591: gap of unknown length
93592 99952: contig of 6361 bp in length
99953 100052: gap of unknown length
100053 107237: contig of 7185 bp in length
107238 107337: gap of unknown length
107339 113752: contig of 6415 bp in length
113753 113852: gap of unknown length
113853 119355: contig of 5503 bp in length
119356 119455: gap of unknown length
119456 127315: contig of 7860 bp in length
127316 127415: gap of unknown length
127416 137358: contig of 9943 bp in length
137359 137458: gap of unknown length
137459 148311: contig of 10853 bp in length
148312 148411: gap of unknown length
148412 160062: contig of 11651 bp in length
160063 160162: gap of unknown length
160163 179879: contig of 19717 bp in length.

FEATURES
source

1..179879
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-799L19"
1..1496
/note="assembly_name:Contig7"

misc_feature

misc_feature 1597..2949
/note="assembly_name:Contig9"
misc_feature 3050..4409
/note="assembly_name:Contig10"
misc_feature 4510..6137
/note="assembly_name:Contig11"
misc_feature 6238..7284
/note="assembly_name:Contig12
clone_end:SP6
vector_side:right"
misc_feature 7385..9013
/note="assembly_name:Contig13"
misc_feature 9114..10444
/note="assembly_name:Contig14"
misc_feature 10545..12153
/note="assembly_name:Contig15"
misc_feature 12254..14386
/note="assembly_name:Contig16"
misc_feature 14487..16990
/note="assembly_name:Contig17"
misc_feature 17091..19022
/note="assembly_name:Contig18"
misc_feature 19123..20497
/note="assembly_name:Contig19"
misc_feature 20598..22209
/note="assembly_name:Contig20"
misc_feature 22310..24677
/note="assembly_name:Contig21"
misc_feature 24778..27655
/note="assembly_name:Contig22"
misc_feature 27756..30189
/note="assembly_name:Contig23"
misc_feature 30290..33658
/note="assembly_name:Contig24"
misc_feature 33759..36912
/note="assembly_name:Contig25"
misc_feature 37013..41069
/note="assembly_name:Contig26"
misc_feature 41170..43658
/note="assembly_name:Contig27"
misc_feature 43759..46574
/note="assembly_name:Contig28"
misc_feature 46675..49890
/note="assembly_name:Contig29"
misc_feature 49991..53751
/note="assembly_name:Contig30"

Query Match 36.6%; Score 1448.2; DB 2; Length 179879;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1456; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 3 GAATTCATCCATTTAAATCATACAAATTAATGCGTTTATAGTATATATACACAGGTTGTCAT 62
Db 109955 GAATTCATCCATTTAAATCATACAAATTAATGCGTTTATAGTATATATACACAGGTTGTCAT 110014
Qy 63 CCATCAATCCATTTTAAACAGTTTATATCTCCAAAATAAACCCCTGCATTCCTTAG 122
Db 110015 CCATCAATCCATTTTAAACAGTTTATATCTCCAAAATAAACCCCTGCATTCCTTAG 110074
Qy 123 CCATCACCCCCCAACATCCCTCCATCCCTCCCTTCCAAAGCCCTGGGCAACCAACATCTACTT 182
Db 110075 CCATCACCCCCCAACATCCCTCCATCCCTCCCTTCCAAAGCCCTGGGCAACCAACATCTACTT 110134
Qy 183 TCTGCTCTATAAATTTGGCAATTCCTGACATTCATATAAATGAAGCAACCAACATGT 242
Db 110135 TCTGCTCTATAAATTTGGCAATTCCTGACATTCATATAAATGAAGCAACCAACATGT 110194
Qy 243 GAGACTTTGACCTGGCTGCTTTCCACTTAGCATTTCTATTTTAAAGGCTCATTTATGTTACA 302
Db 110195 GAGACTTTGACCTGGCTGCTTTCCACTTAGCATTTCTATTTTAAAGGCTCATTTATGTTACA 110254
Qy 303 GTACTAGCAGTACTTCACTCTCTTTTATTTCTCAATGGTATTCACGTGTGGGTATCC 362

Db 110255 GTACTTAGCAGTCTCATCTCTTTTAACTCTCAAAATGGTATTCCAGTGTGTGGGTATCC 110314
Qy 363 CATATCATATTAATAGACAGAGTCTCAGTCTGTCCACCCAGGCTGGAGTGCAGTGGCAC 422
Db 110315 CATATCATATTAATAGACAGAGTCTCAGTCTGTCCACCCAGGCTGGAGTGCAGTGGCAC 110374
Qy 423 AATCATAGCTCACTGTGAACCTCAAACTCCTGGGCTCAAGTGATCTACTACCTCAGCCTC 482
Db 110375 AATCATAGCTCACTGTGAACCTCAAACTCCTGGGCTCAAGTGATCTACTACCTCAGCCTC 110434
Qy 483 CAGAGTAGCTAGGACTACAGGCACACAGGCATACCTGGCTGAATTTTTTTTTTAATTT 542
Db 110435 CAGAGTAGCTAGGACTACAGGCACACAGGCATACCTGGCTGAATTTTTTTTTTAATTT 110494
Qy 543 TCATTTATGATTAATTAATTTCTTTCTTTTGTGTTGTTGTTGTTGTTGATAGGTTCTCAC 602
Db 110495 TCATTTATGATTAATTAATTTCTTTCTTTTGTGTTGTTGTTGTTGATAGGTTCTCAC 110554
Qy 603 TTTGTTTACCAGGCTGGAGGGCAGTGGCATGGTGACAGCTGAGCAGCCTTGACTTCCTGG 662
Db 110555 TTTGTTTACCAGGCTGGAGGGCAGTGGCATGGTGACAGCTGAGCAGCCTTGACTTCCTGG 110614
Qy 663 GCTCAAGTGATCTCTCCTCCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGCTGTACCC 722
Db 110615 GCTCAAGTGATCTCTCCTCCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGCTGTACCC 110674
Qy 723 ATGCTGCTGATATTTTTTTTCTTGAACAGGATACACTGTGTTGCCAGGCTGGAGT 782
Db 110675 ATGCTGCTGATATTTTTTTTCTTGAACAGGATACACTGTGTTGCCAGGCTGGAGT 110734
Qy 783 ACAGTGGGCTAATAATAGCTCACTGAGCCTCCCTCCTGGGCTCAAGCAATCCGCTGGC 842
Db 110735 ACAGTGGGCTAATAATAGCTCACTGAGCCTCCCTCCTGGGCTCAAGCAATCCGCTGGC 110794
Qy 843 CTCAGATCTGAGTAGCTGGGACTACAGGCTGTGTCACACAGGCCAGCTAAGTTTAA 902
Db 110795 CTCAGATCTGAGTAGCTGGGACTACAGGCTGTGTCACACAGGCCAGCTAAGTTTAA 110854
Qy 903 AAAATGATTTTGGTATAGAGAGGCTCTGCTATGCTGCTAGGCTGATTTTTTATTTGTT 962
Db 110855 AAAATGATTTTGGTATAGAGAGGCTCTGCTATGCTGCTAGGCTGATTTTTTATTTGTT 110914
Qy 963 GAGCAAGGCTCTCACTATGTTGCCATGATPCCGCCACCTCCACTTCCCAAGAGTCTATC 1022
Db 110915 GAGCAAGGCTCTCACTATGTTGCCATGATPCCGCCACCTCCACTTCCCAAGAGTCTATC 110974
Qy 1023 TTATCTGTTTATTAAGTATGACAGACATTTAGTGTGTTTCCACTTTTGACCAATATG 1082
Db 110975 TTATCTGTTTATTAAGTATGACAGACATTTAGTGTGTTTCCACTTTTGACCAATATG 111034
Qy 1083 AATAATCTCCAGTGAATATTCATGTATACATTTGTTGGGCATATGTTTTCATTTCTGT 1142
Db 11035 AATAATCTCCAGTGAATATTCATGTATACATTTGTTGGGCATATGTTTTCATTTCTGT 111094
Qy 1143 TGGGTTTATATCTAGAGTGGAAATGCTGGATCCCGGTPAATTTTGAAGGAGAGTT 1202
Db 111095 TGGGTTTATATCTAGAGTGGAAATGCTGGATCCCGGTPAATTTTGAAGGAGAGTT 111154
Qy 1203 CAGGGAAGAAACCTTGGGAAAATGAACATGTTTAGAAATCAGCAAGAGTGCAGGGT 1262
Db 111155 CAGGGAAGAAACCTTGGGAAAATGAACATGTTTAGAAATCAGCAAGAGTGCAGGGT 111214
Qy 1263 TTTTCCGAGTTTATTTTATTTCTGTTGACAAATGTCAGTTTGTATGAAGATACAAGTT 1322
Db 111215 TTTTCCGAGTTTATTTTATTTCTGTTGACAAATGTCAGTTTGTATGAAGATACAAGTT 111274
Qy 1323 ATACTAAGTGAGAGTGAGAAATTAAGGCTGGAATAGGCGCTTCAAGTAAATCATGAAG 1382
Db 111275 ATACTAAGTGAGAGTGAGAAATTAAGGCTGGAATAGGCGCTTCAAGTAAATCATGAAG 111334
Qy 1383 CACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAACAAAATATAAATAAATAATCAAAAT 1442
Db 111335 CACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAACAAAATATAAATAAATAATCAAAAT 111394

Qy 1443 TTTTCTTTTCTTTTGGAGAAAGAGTCTTGC 1471
Db 111395 TTTTCTTTTCTTTTGGAGAAAGAGTCTTGC 111423

RESULT 9
AC139669
LOCUS
DEFINITION
AC139669
AC139669.1 GI:28273352
VERSION
HTG; HTGS PHASE1; HTGS DRAFT.
KEYWORDS
Lemur catta (ring-tailed lemur).
SOURCE
ORGANISM
Lemur catta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Strepsirhini; Lemuridae; Lemur.
1 (Bases 1 to 190784)
Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
Peng,Z., Malinov,I. and Rubin,E.M.
Direct Submission
Unpublished
2 (Bases 1 to 190784)
Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
Peng,Z., Malinov,I. and Rubin,E.M.
Direct Submission
Submitted (08-FEB-2003) Genome Sciences, Lawrence Berkeley National
Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
Sequence Produced by Berkeley PGA
Web site: <http://pga.lbl.gov>
Center Code: PGABERK
Center Project Name: L110
Bac Clone Name: LB2-62P23

This sequence has been compared to sequences of other species
using Vista (<http://www-gsd.lbl.gov/VISTA>). The results can be
viewed at:
http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=SCD

The order-orientation of the draft sequence was accomplished by
using:

Avid (<http://baboon.math.berkeley.edu/mauid>),
Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

If the Bac Library Name is LB1 to LB4, please see website
for the description: <http://www-gsd.lbl.gov/cheng/BAC.html>
These libraries are available through the BACPAC Resources Center:
<http://www.chori.org/bacpac/libraryres.htm> as LBNU-1 to LBNU-4.

Summary Statistics:

Sequencing vector: Plasmid; pUC18

Chemistry: Dye-terminator Big Dye

Assembly program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently
consists of 1 contigs. The true order of the pieces

* is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence.

* as soon as it is available and the accession number will
be preserved.

* 1 190784: contig of 190784 bp in length.

Location/Qualifiers

1. .190784

/organism="Lemur catta"

/mol_type="genomic DNA"

/db_xref="taxon:9447"

/clone="LB2-62P23"

FEATURES
source

ORIGIN

Query Match	20.6%; Score 816.5; DB 2; Length 190784;
Best Local Similarity	73.7%; Pred. No. 2.7e-189; Indels 106; Gaps 23;
Matches 1386; Conservative	0; Mismatches 389;
2124 AGCGAGATCTCTCTCAAAAAAACAAGAAATTAAGCAATTAAGCATGTCAG	2183
120071 AGCGAAATGGACTAAGACAAAGAGTAACCCCTGCAGTTGGTAAGACCAAGATCTTCCAA	120130
2184 AGAAGAACTGAAGGGGTGAGACACAGTACAGATTTCTGTGCCACATGCAAGTACTTCT	2243
120131 AGCAAAAGCTGAAGGAAGCAGCATGTGCTTGGCTTCTAGTCCAGCAGCAAGCATTTCT	120190
2244 GAGGCATGACTGGATGAGCTGTCCACATCTGAATATCATCCAGTCTTGTTCAGAACTTTCA	2303
120191 GAGTCAATTTAGATGAGCTGTCCATATCTGAACATCATAGTCCCATTCAGAACTTTCA	120250
2304 CACCGACAGGAGGAGCAGACATGGAATAGTCTCTCTGTCTACCTGGCCAGAGATTTGCC	2363
120251 CACTGACAGAAAGCCAGACAGGAATAGACATCTGGTCACTGGCCAGAAAGTTGTCC	120310
2364 TTGACCTTGAGACCTGAGCCACAAAGAGAGCTGTAGTCTTACCTCCAGCAAAAT-CCC	2422
120311 -TGACCCAGAGACACAGCCGCAACAGGAGCTGTAGTCTCTGTCCAGGCATCCCC	120369
2423 AGGTGCTTGTCTTCTTGGGAAGTGAATCAATGGCGCAGCCTCCGTATTTTCTCTCTTTC	2482
120370 AAGTGTCTCTCTCTCTGAGACAGATGGTGGCAAGAACTTTCTTTTGTCTCTTTC	120429
2483 CCAGGGAAGGATTCCTAGGCGAGTATTTGGGAAGACATGGGCATGGAAGACACCGGGT	2542
120430 CCAGAAAGGATCTCTGGGCGAGTATTTGGG---GAGACAGCAGGAGGAGAGCTGGGT	120486
2543 GAATGATAGCTGCTGCTTCTGAGCTCTCATGTGTAAGGCTCTTACAGACACGGAAG	2602
120487 GAATACAGAGCTGCTGCTTCTGCTGTACGCTGAGGCTCCACAGGCCAGAAAT	120546
2603 ATGGGGGCAACAGGACAGATCAGTAGGGTCAGAGCATCTCAGGGACAGGAGGAGCTGGG	2662
120547 ATAGGGGCAAGAG-----AGACAGGTGACAGCATCTCAGGGACCTAGGGCC-----	120591
2663 TCTGACAGGAGTAAAGACTTGGGCTCTCATATGGTGTCTGGCTCAACTGCCAGC	2722
120592 TTCTGAGATGGAGTAAGAGCGCGCGGTGTCATTAACATTTCTGG---CAATTCCTAGC	120649
2723 TCCGTCACTTACTGCTGTGTGACCATGGGCAAGTATTCATCTCTCCATATCT--CT	2780
120650 TTACACACTTACCGTCTGCTGTGACCATTTGGCAAGTTATTTACCTCTCCGTATCTCACT	120709
2781 TTCTCTACTTTTAAATGGAATTAATGGGGTACCACCTCCAGGCTCAGAGAGGCTTA	2840
120710 TTCTCTACTTTTAAATGGAATTAATGGGGTACCACCTTCAATAGGTCAAGTGAGGATTT	120768
2841 CAGAAACAGATCTCTGGAATTCGCTTGCAAGTAATTAATCAATPACCTGCGAGTATTTCT	2900
120769 ----AAATGATCTCTGGAATTTGCTCGAGTAACATTAATCTGCTAATCTCTT	120824
2901 ATTCACATCCAGACCTTTTCCGCTGCTGTGGTGAAGCAACATGTCTGTCTTCTGA	2960
120825 -TTCCACTACAGTCTTTTGGCTGTATCTGGGTGAAGCAACAGTCTGTGTCTCTCTG	120883
2961 CGGTTTCCAGAAAGATTTCCAAATTAACAACCTG-CCAGTCTGAAGATCTCCAAAC	3019
120884 TTTTTCATTAAGA--GTTCAAAATTAACAACCTGCAATCTGAAGATCTCCAAAC	120941
3020 ATCCGACAGCATCTCTGAGCGCGGCTTGGGATGGGACTCCCGCCCGGGTCTCTGAA	3079
120942 GTCCCGACAGCTTCTAGAGCGCTGGGCAAGGGGTGGTCTGTGGCCACTGTGCTGGG	121001
3080 CAGGA-----TGCGTGGCGGAGCAGACAC	3105
121002 GTGAGAGTCCCTGCGGAGGTACCCAGTGGGGGCAACAGGTGGCGGAGCAGAC	121061
3105 GTGAGAGTCCCTGCGGAGGTACCCAGTGGGGGCAACAGGTGGCGGAGCAGAC	3159


```

REFERENCE
AUTHORS      1 (bases 1 to 17088)
TITLE        Medrano,J.F., Islas-Trejo,A.D. and Johnson,A.M.
JOURNAL      Genomic sequence of the bovine stearyl-CoA desaturase
REFERENCE    2 (bases 1 to 17088)
AUTHORS      Medrano,J.F., Islas-Trejo,A.D. and Johnson,A.M.
TITLE        Direct Submission
JOURNAL      Submitted (21-FEB-2003) Animal Science, University of California,
              One Shields Avenue, Davis, CA 95616-8521, USA
FEATURES
source
1..17088
    /organism="Bos taurus"
    /mol_type="genomic DNA"
    /db_xref="taxon:9913"
    /chromosome="26"
1..17088
    /gene="SCD"
1..1880
    /gene="SCD"
1..1880
    /gene="SCD"
join(1881..2051,2568..2850,6720..6850,8314..8519,
10099..10331,13019..17088)
    /gene="SCD"
/product="stearyl-CoA desaturase variant A"
join(2025..2051,2568..2850,6720..6850,8314..8519,
10099..10331,13019..13218)
    /gene="SCD"
/EC number="1.14.99.5"
/note="delta-9 desaturase"
/codon_start=1
/product="stearyl-CoA desaturase variant A"
/protein_id="AA063569.1"
/db_xref="GI:29469127"
/translation="MPALLQEEISSYTTTITAPSPRYLQNGGKLETPLYLEE
DIRPEMRDDIYDPTQDEGPKLEYVVRNIIIMSLHLGALXGIIIPCKLYTYI
WLFYLMGALGITAGARHLWSHRTYKARLPLRVFLITGNWAFQNDVFEWSRDRHRAH
KVFSTADADPNRSRGFFSHVGMILLVKKHPAVKEKSTLNLSDLRAEKLVMFQRRYY
KPGVLLCFILPTLVPWLWDETQNSLFFATLFRYALGLNVTWLVNSAAHMYGYRYP
KGTINPRNIIIVSLGAAGEGPHNYHHTFPYDSAYSEYWHINFTTFIDCMAAIGLAY
DRKNVSKAAIIRIKRTGEESYKSG"
6926
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="g"
8243
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="g"
8646
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="g"
10153
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="g"
10213
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="c"
10329
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="t"
13565
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="t"
13638
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="c"
13757
    /gene="SCD"

```

```

variation
    /note="compared to sequence in variant B"
    /replace="a"
14047
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="t"
14578
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="g"
14662
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="y"
15001
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="g"
15101
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="t"
15202
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="a"
15284
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="c"
15491
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="a"
16524
    /gene="SCD"
    /note="compared to sequence in variant B"
    /replace="t"

```

ORIGIN

```

Query Match      13.2%; Score 521; DB 4; Length 17088;
Best Local Similarity 64.9%; Pred. No. 7.7e-117; Gaps 24;
Matches 1232; Conservative 0; Mismatches 490; Indels 177;
QY 2164 AAGCAATTAGACATTGCAGAGAGAACTGAGGGGGTCAGACACGTCACAGATTCTGT 2223
Db 1 AAGCCATTAGACATTGCAG-GAGAGCTTGAGAGAGCAACACCTGCCAGACTTCTCT 59
QY 2224 GCC----ACATGCCAAAGTACTTCTGAGGCATGACTGGATGACTGTCCACATCTGAATC 2279
Db 60 CCCACAAACATGCTGAGGGCTTCTGAGCGGTGA-TGAGACAGCTCTGTATATATAAAC 118
QY 2280 ATCAGTCTTGTTCAGAACTTTCACACGGACAGGAGCCAGACTGGAATCAGTCTCC 2339
Db 119 ATCTAGTCTCATTCATCAAACTTTTGACCTGAGAGAGTGTGAGACTGATGATGAGGTGA 178
QY 2340 TGGTCACTGGCCAGAGAGTTGGCCCTTGACCTGAGACCACTGGCCCAACAAAGAGTGTCT 2399
Db 179 TGGTACCGGCCAGAGAGCTGGCC-AGACCCCAAGACCTGTGGCCTTCTAGGAGTCTCT 237
QY 2400 TAGTCTACTCCCGAGAAATCCAGGTGCTTGTCTTCTTGGGAAGTGAATCATTTGGCGCA 2459
Db 238 TAATCATCTCCAGGCAATCCAGTGCTCTCTTCC--AGAGACTAGTGTGTCGACCA 295
QY 2460 GCATCCGATTTTCTCTCTTCCAGGGGAGGATCTTAGGGCAGTATTTGGGAAGAC 2519
Db 296 AAATCTCTCACTTCTTCTCTTCCCA-----AATGATGATATTTGGGAGAC 343
QY 2520 ATGGGCATGGAAGGACACCGGGTGAATGATCATAGCTGCTGTGTTCTAGCT----CTCAT 2575
Db 344 AGGGGTGTGTTGGAGCCCGGAGAGATGCACAGCTGCTGCTGTTCCAGCTTATCCCAT 403
QY 2576 GGTAAAGGCTCTTACACACACCGGAAAGATGGGGGCACAGGACAGATCATGATGGGTGAGA 2635

```

Db 404 GGTAGGCTCCACACGCCCGGAAAGGCAAGACTGTGAGCCACAACAGTAGGTCAGA 463
 QY 2636 GCATCTCAGGAGCCGAGGCGCAATATGCTCTGAGCAGGAGTAAGAGCTTGGGCTCTCAT 2695
 Db 464 ACATCTCAGGAGCCAGG-----GATCTGAGCAGGAGTAAGA-CTTGGGCTCCATC 514
 QY 2696 ATGGTGTCTTCTGGGCTCAACTGCCAGCTCGTCACTTATCTGTTGCTGTGACATGG-GC 2754
 Db 515 ATTACATTTCTGAGCTCAGCTACTAGCTCCATCATTTGATTTGGCTGTGTGACATAAAGC 574
 QY 2755 AGTTATTCATCTCTCCATATCTCTTCTC---ACTTTTAAATGGAATATGGGTGA 2811
 Db 575 AGTTATTTAAATTTCTCCATAGTCACTCTCTCATCTTTTAAAGGAAATAATGACTA 634
 QY 2812 CCACCTCCAGAGGTGACA--GAGAGCTTTACAGAAACGATTTCTTGTGAATTTGGTTCG 2869
 Db 635 CTTACCTCAGAGGCGACAAGTTAAGATCAATGAATGATTTCTTGTGAATTTGGCTTG- 693
 QY 2870 AGTAATAAATCAATACCTGCCAGCTATCTTATTTCCATCTCAAGCCCTTTCGCTGCTG 2929
 Db 694 -----CAATACCTTTTGGCTACTCTTGTGTGCACTTACAGTACTTTCTCTGCTT 743
 QY 2930 CTGGGTGAACACATCTCAGTGTTC-----TGACGGTTTCCACAAAGAAATTC 2981
 Db 744 CTGAGTGAAGACAGCTCTGCTCTTTCTCCAGAAATGCGCTTTCCCAACGAGCAATTC 803
 QY 2982 CAAATTTACAACTGCCAGCTGTAAGAAATCTCCAAACATCCCGCAGCATCTCTGAGGC 3041
 Db 804 CAAACACAGCCGAGCTTCT-----AATCTCCAAACATCCCGCAGCATCTCTGTTGGC 859
 QY 3042 GCGGGCTTGGGATGGGACTGCCCGCGCGGTC----- 3074
 Db 860 ATGGGCAAGGACAGCAGCAGCTCACTGGTGGCTGGCGGTGAACATTTCTTTGGCGGTG 919
 QY 3075 -----CTGAACAGGATGGTGGCGGAGGACACACACAGCAGCAGCTCTGTGTG 3126
 Db 920 ACATTTGGCTGAACCTCTTACGCGGTGCAGATCTACACATCAGCGCATCTACTTGTGTG 979
 QY 3127 CGGCGGAGTCCGG-----TGCGGTCCG 3150
 Db 980 TGCGGCGCTCCAGCTGCGCTGCGGAGGAGGAGCCGCAACCTGTTACAGCATGCCG 1039
 QY 3151 GGTGAGCAGCGGTGGCTGGTGGCGGAGAGCCATTTGTCAGCGTACCGA----- 3206
 Db 1040 GGTAGCGCTCGCGCTGGTGGCGGCGGCGGCGGCAATGTTTCGCGCGGAGCAGAACCG 1099
 QY 3207 -----GCCCCCGCGCTCGCCCGGAGGAGGCGGCGCTTCCC 3244
 Db 1100 TCCCTGCGCCCTACCCCGCATCCCGCGGCTCGGACAGAGGAGGCGGCGCTTGGC 1159
 QY 3245 CGTCCCAAGCTCCAGATCTGGGCTGGTGCACGCTCTCCCTGCGCAGCGGCTTGGGG 3304
 Db 1160 CGGAGCGCGGCCCGGCGCGGCTTTGAGGACAGCTCTCCCTCTCTACGCTCTAGACG 1219
 QY 3305 GACGGGAGACGGGACGAGATGTTAGTGGTGGCGGCCCGGAGGTTTACCACTGTTT 3364
 Db 1220 AGCGGAAAGGCTGGGGTGGAGTGTGTGAGTGTGCGCTCGTGGTTTCCCACTCTTT 1279
 QY 3365 CTTGAGAACTTCCCGAGTCCCAACCCCTTCTCGTGTGCGCGGAGGCGGCTGCTG 3424
 Db 1280 CCCAGGAACTTCCCGAGTGCCTTCCCA-----TTTGGAAATTTGCGGCGGCGCTGCTG 1334
 QY 3425 GGCT-----AGGCTCCGCGCCAGCGCCCAACCGGGTCCCGAGCCCTTCCAGAGAGAA 3480
 Db 1335 GGCTGGCAGCATCCCGGCGCACTCCCGAGTGGGTCTCTCTCTCTCCCGAGCGCT 1394
 QY 3481 GCTCCCGAGCGGGATGCGGCGAGGCGCCAGCGCGGTGAGAGAGGCTGAGAGG 3540
 Db 1395 CAGAGCG--GCAAGGTTGCGGTGAGGCGCCAGCGGCCATGTAGAGAGCGG---AGG 1449
 QY 3541 AGAACAAGAGGAGGAGGAGGAGGAGCTGGCGGAGAGGAGCAAGCAGATTGGCGCGA 3600

Db 1450 AGAAGGAGGAGGAGGAGGCTAGTAGGAGCTCGCGGAGAGGAGAAACAGCAGATTGGCCGA 1509
 QY 3601 GCCATGCGACGCGCAGGAGGAGTGGACCAAAATTCCTTCGCGCAATGACAGCCGGA 3660
 Db 1510 GCCAATGGCAACGCGGAGGAGGAGTGGACCAAAATTCCTTCGCGCAATGACAGCCGGA 1569
 QY 3661 GTTTACAGAAAGCTCATTTAGCAATTTCCCGAGAGGAGGCGGAGGAGGAGGCGGCTG 3720
 Db 1570 GTCTACAGAGCC-CATTAGCAATTTCCCGAGGAGGAGGAGGAGGAGGAGGCGGCTG 1628
 QY 3721 TGTGTTGTCGCTGTCGAGCATCCCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 3780
 Db 1629 GCCAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1686
 QY 3781 TCTGCTCTCTCCCGCTCCCGCTTACCTCCAGCGGAGCCCGCGCGGAGTCAACTCC 3840
 Db 1687 TCGGCGCTCTGCT 1746
 QY 3841 TCGCACTTTGCCCT 3900
 Db 1747 GCGCACTTTGCCCT 1806
 QY 3901 CACCGTTGCCAGCTTACGCTTTAAATTTCCCGCTCGG 3939
 Db 1807 CACCCCTGCCAGCTAGCTTTAAATTTCCCGAGCAG 1845

RESULT 11
 AX357255
 LOCUS 617 bp DNA linear PAT 13-FEB-2002
 DEFINITION Sequence 1 from Patent WO0162954.
 ACCESSION AX357255
 VERSION AX357255.1 GI:18674409
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Theria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Brownlie, A.J., Hayden, M.R., Attie, A.D., Ntambi, J.M.,
 Gray-Keller, M.P. and Miyazaki, M.
 TITILE Methods and compositions using stearyl-coa desaturase to identify
 triglyceride reducing therapeutic agents
 JOURNAL Patent: WO 0162954-A 1 30-AUG-2001;
 Xenon Genetics Inc. (CA); WISCONSIN ALUMNI RESEARCH FOUNDATION
 (US); University of British Columbia (CA)

FEATURES
 source
 1..617
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 12.6%; Score 498.4; DB 6; Length 617;
 Best Local Similarity 99.8%; Pred. No. 1.6e-111;
 Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3459 CCCAGCCCTTCCAGAGAGAAAGCTCCCGACGCGGATGCCGCGGAGAGGCCCGCGCG 3518
 Db 4 CCCGCCCCCTCCAGAGAGAAAGCTCCCGACGCGGATGCCGCGGAGAGGCCCGCGCG 63
 QY 3519 GGTGGAAGAGAGCTGAGAGAGAAACAGAGGGGAGGGGAGGAGGCTGGCGGCGAG 3578
 Db 64 GGTGGAAGAGAGCTGAGAGAGAAACAGAGGGGAGGGGAGGAGGAGGCTGGCGGCGAG 123
 QY 3579 AGGGAACAGCAGATGCGCCGAGCAATGCAACGCGAGGAGGAGGAGGAGGAGGAGGAGG 3638
 Db 124 AGGGAACAGCAGATGCGCCGAGCAATGCAACGCGAGGAGGAGGAGGAGGAGGAGGAGG 183
 QY 3639 CTTTCGGCCAAATGACAGCGCGGAGTTTACAGAGCCTCATTTAGCATTTCCCGAGGCGAGG 3698
 Db 184 CTTTCGGCCAAATGACAGCGCGGAGTTTACAGAGCCTCATTTAGCATTTCCCGAGGCGAGG 243

	QY	3699	GGCAGGCGGCAGACGCCGGGTGTGTGGTTCCTCCCGCCCCTTTACTCATCACGCGGG	3818
	Dd	244	GCCAGGCGGCAGACGCCGGGTGTGTGGTTCCTCCCGCCCCTTTACTCATCACGCGGG	363
	QY	3759	GCGGTTCGCGCGAGCCTTCGGCTTGCTCTCCCTCCCGCCCCTTTACTCATCACGCGGG	3938
	Dd	304	GCGGTTCGCGCGAGCCTTCGGCTTGCTCTCCCTCCCGCCCCTTTACTCATCACGCGGG	483
	QY	3819	ACC GCC CCG CGCC AG TCA ACT TC CT CG CA CT TT GC CC CCT TG CG A G GG AT AA A A G CG GG	3878
	Dd	364	ACC GCC CCG CGCC AG TCA ACT TC CT CG CA CT TT GC CC CCT TG CG A G GG AT AA A A G CG GG	423
	QY	3879	CTGAGGAATACCGGACAACGCTTCACCGTTCAGCTTTAAATTCCCGGCTCG	3938
	Dd	424	CTGAGGAATACCGGACAACGCTTCACCGTTCAGCTTTAAATTCCCGGCTCG	483
	QY	3939	GCGACCTCCACGACCGCGG	3958
	Dd	484	GCGACCTCCACGACCGCGG	503
RESULT	LOCUS	AL157935	Human DNA sequence from clone RP11-203J24 on chromosome 9, complete sequence.	PRI 09-JAN-2002
	DEFINITION	AL157935		
	VERSION	AL157935.28	GI:17221172	
	KEYWORDS	HIG.		
	SOURCE	Homo sapiens (human)		
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. Corby,N.		
REFERENCE	AUTHORS	Submitted (09-JAN-2002) Wellcome Trust Sanger Institute, Hinxtom, Cambridge@shire, CB10 1SA, UK. E-mail enquiries: hamquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk		
TITLE	JOURNAL	On Nov 30, 2001 this sequence version replaced gi:16944853. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.		
COMMENT		This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EM:, EMBL; SW:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep/ This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr9 RP11-203J24 is from the library RPT-11.1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pBacc3.6		
FEATURES	source		The entire insert of clone RP11-203J24 The true left end of clone RP11-379C10 is at 163899 in this sequence. The true right end of clone RP11-228B15 is at 66809 in this sequence.	
			Location/Qualifiers 1..164201 /mol_type="genomic DNA"	

D	b	5788	CAGCCAGGGCTGGGATTCGTGCCCAAGTAGGCCTGTGCTCCACATTTTCTTTCTTTT	5847
Q	y	1191	ACAGGCAGAGCTTCAGGGGAAGAAAACTTGGGAAAAATGAAGCAATGTTTAAAGAAATCAGCAA	1250
D	b	5848	TCACTTTTCTTTTCTGAGATGGAGTTTCACTCTT---CCTGCCAGGCTGGAGTCG	5904
Q	y	1251	GAGTGCAGGGTTTTTTCGGAGTTTTATTTATATTCGTGTGA CAATGTGCAAGTTTGATG	1310
D	b	5905	GAGACCAGCCTCATCAACATGGTGAAACCCCAATGGTCCGATCCTAGCTCACTGTGTTCTT	5964
Q	y	1311	AAGATACAAGTTATCTAAGTGAAGAGTGAATTAAGCTGGAATAGGGCGTTCAGACT	1370
D	b	5965	TTTTTTAGGCTCGT--CAAGTGAAGCAATCAGAGTGGAGAAGGAACAAGCAATCTGTAA	6022
Q	y	1371	AAAATCATGAAGCACTTTGAATCAAAAAATTAAGGAGCTTGGCTGTAAAACAAAATAATAA	1430
D	b	6023	CTAGTCTGTGATCAATTACTTTGTAACACAATGTTTTTTCATTTAGTGTGCCATTTAG	6082
Q	y	1431	AAAATCAAAATTTTTTTTTTTTTTTTGGAGAAAGATCTTCTCTTTCAOCTGGCTGGAG	1490
D	b	6083	TTTTTTTCTATTATTAATTTTTTTCAGATGGAGTCTTGTCTGTGT--TGCCAGGCTGGAA	6141
Q	y	1491	GGCAGTGGTGTGATCTCAGCTCACTGCAACTTTGCCTCCCGGGTTCAGCAATTTCTCCT	1550
D	b	6142	TGCAATGGCGGATCTCGGCTCACTCGGCACTCCACCTCTCTGGGTTCAAGTGATTTCTCT	6201
Q	y	1551	GCTTCAGCTCCCAAGTAGCTGGGACTACAGGCACTTCCACCAATGCCCAAGCTGATTTTT	1610
D	b	6202	GCTTAGGCTCCCAAGTAGCTGGGACTTACAGTGCTATGCCACATGCCCAAGCTAATTTT	6261
Q	y	1611	GTAATTTTAGTAGAGATGGGATTTTCACTTTGTTGGCCAAGCTGGTCTCAAACT--	1663
D	b	6262	GTAATTTTAGTAAGATGGGGTTTGGCAATGTTGGCCAGGATGGCTCAAACTCCTGACC	6321
Q	y	1664	----TTTTGCTGCATAATTTGTTAATATTTCTTTTCTGAGGTAGGCGCCCCAG	1719
D	b	6322	TCAGTGATCCACCGCCTTGGCCTCCCAAGTCTAGGATATAGCGGTGAGCACTGC	6381
Q	y	1720	ACCAAAAAAATAATCTTAGAATCCAAATCAGGTGTG--GTTTGACCA	1768
D	b	6382	ACCGGCGCTCTCATAATTTTCCACATTTTGTGTAGCAATCCTTCCTAGCAAT	6441
Q	y	1769	CTGTCACTTGAGAACCA CAGTGTGACACCGGCTCAGGAGTAGAGTGATCTCGTCTGA	1828
D	b	6442	TTTTCTAACCTTATTCOCAAGAGAAGAAACACAGCAATCTACAAGCCAGTGTCTATT	6501
Q	y	1829	AAGAGAAATAGATGAATAATTTCT-----TTTTTTTTTTTTTTTTTTTTTTTTTT	1853
D	b	6502	GAAAGATGACCTCAAGACTTTCTGAATAGTGGCATGTGAACAGGTGAGTGTGTAGAC	6561
Q	y	1854	-----CCGGGCCAGCGCTGGTGGCTCATGCTGTAAATCCACGACTTTTGGGAGGCCA	1905
D	b	6562	TAAGACAGGTAGACCGGCTGGTGGCTTA CGCTGTAAATCCAGCACTTTGGGAGCTG	6621
Q	y	1906	AGGCATGTGGATCACTGAGSTCAGAGTTCAAAACCAAGCTGGCCAAACATGTGTAACC	1965
D	b	6622	AGGTGAGCCGATCACTGAGGTTGGGAGTTCAAGACCAACCTGACCAACGTGGAGAAAC	6681
Q	y	1966	CGGTCTCTACTAAAAATACAAAAAATTAGCTAAGTGTGGTGGCCATGCCCTGTAATCCCA	2025
D	b	6682	CGGTCTCTACTAAAAATACTA--AATTAGCCAGGTGTGGTGGCGCATGCCCTGTAATCCCA	6739
Q	y	2026	GCTACTTGGGAGGTTGAGGCAGGAGAAATTTCTTGAA--CCCGGAGGCAGAGTTGCAATG	2084
D	b	6740	GCTACTCGGGAGGCTGAGGCAGAGAGATCGTTGAAACCCAGGAGCGAGGTTGTGTG	6799
Q	y	2085	AAGCGAGATCACACCACTGCACTCCAGCGCTGG--GGAGAGAGCGAGATCTTCTCTCAAAA	2143
D	b	6800	AGCCAAGATCGCGCCATTGAGCTCTAGCTTGGCTGGGAGCAAGAGTGAATCTCGCCTCAAAA	6859
Q	y	2144	AAACAAAAAACAAAGAA	2161
D	b	6860	AAAAAAAAAAAAAAAAA	6877

AC119725	20288 bp	DNA	linear	HTG 27-MAR-2003			
LOCUS	Homo sapiens chromosome 3 clone RP11-759D18, WORKING DRAFT						
DEFINITION	SEQUENCE, 36 unordered pieces.						
AC119725							
AC119725.2	GI:29239397						
HTG; HTGS_PHASE1; HTGS_DRAFT.							
KEYWORDS	Homo sapiens (human)						
SOURCE	Homo sapiens						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
AUTHORS	1 (bases 1 to 200288)						
	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbarga,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinu,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C., Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Loulssed,H., Lozard,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Minal,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokweso,S., Ogwu,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sonaika,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.						
	Direct Submission						
TITLE							

COMMENT

Web site: <http://www.hgsc.bcm.tmc.edu/>
 Drafting Center Code: BCM
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: HCR
 Center clone name: RP11-759D18
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 169013 bases at least Q40
 Consensus quality: 174569 bases at least Q30
 Consensus quality: 178019 bases at least Q20
 Estimated insert size: 177488; sum-of-contigs estimation
 Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 36 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 1266: contig of 1266 bp in length
 * 1267 1366: gap of unknown length
 * 1367 2448: contig of 1082 bp in length
 * 2448 2548: gap of unknown length
 * 2548 4153: contig of 1605 bp in length
 * 4153 4253: gap of unknown length
 * 4253 5318: contig of 1065 bp in length
 * 5318 5418: gap of unknown length
 * 5418 6987: contig of 1569 bp in length
 * 6987 7088: gap of unknown length
 * 7088 8469: contig of 1282 bp in length
 * 8469 8470: gap of unknown length
 * 8470 9488: contig of 1019 bp in length
 * 9488 9589: gap of unknown length
 * 9589 11671: contig of 2083 bp in length
 * 11671 13295: contig of 1524 bp in length
 * 13295 13396: gap of unknown length
 * 13396 14422: contig of 1027 bp in length
 * 14422 14523: gap of unknown length
 * 14523 15982: contig of 1460 bp in length
 * 15982 16082: gap of unknown length
 * 16082 18344: contig of 2262 bp in length
 * 18344 20093: contig of 1649 bp in length
 * 20093 20193: gap of unknown length
 * 20193 21347: contig of 1154 bp in length
 * 21347 21447: gap of unknown length
 * 21447 24953: contig of 3506 bp in length
 * 24953 25053: gap of unknown length
 * 25053 27672: contig of 2619 bp in length
 * 27672 27772: gap of unknown length
 * 27772 29866: contig of 2094 bp in length
 * 29866 31985: contig of 2019 bp in length
 * 31985 32086: gap of unknown length
 * 32086 34334: contig of 2249 bp in length
 * 34334 37766: contig of 3332 bp in length
 * 37766 37866: gap of unknown length
 * 37866 40539: contig of 2673 bp in length
 * 40539 40639: gap of unknown length
 * 40639 44375: contig of 3736 bp in length
 * 44375 44476: gap of unknown length
 * 44476 50326: contig of 5851 bp in length
 * 50326 50426: gap of unknown length
 * 50426 50427 53739: contig of 3313 bp in length
 * 53739 50427

53740	53839: gap of unknown length				
53840	61829: contig of 7990 bp in length				
61830	61829: gap of unknown length				
61930	68906: contig of 6977 bp in length				
68907	69006: gap of unknown length				
69007	78680: contig of 9674 bp in length				
78681	78780: gap of unknown length				
78781	89431: contig of 10651 bp in length				
89432	89531: gap of unknown length				
89532	100730: contig of 11199 bp in length				
100731	100830: gap of unknown length				
100831	112190: contig of 11360 bp in length				
112191	112230: gap of unknown length				
112291	122606: contig of 10316 bp in length				
122607	122706: gap of unknown length				
122707	135164: contig of 12458 bp in length				
135165	135264: gap of unknown length				
135265	147115: contig of 11851 bp in length				
147116	147215: gap of unknown length				
147216	159369: contig of 12154 bp in length				
159370	159469: gap of unknown length				
159470	173326: contig of 13857 bp in length				
173327	173426: gap of unknown length				
173427	200288: contig of 26862 bp in length.				

FEATURES
 Location/Qualifiers
 1..200288
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="3"
 /clone="RP11-759D18"

ORIGIN

Query Match	11.2%;	Score 447.2;	DB 2;	Length 200288;
Best Local Similarity	58.5%;	Pred. No. 1.6e-98;	Indels 111;	Gaps 14;
Matches 1099;	Conservative 0;	Mismatches 668;		
QY	365	TATCATATATTAGACACAGGTTCTCACTGTGCACCCAGGCTGGAGTGAGTGGCAAA	424	
Db	103464	TTTCTTTCTTTTGGACACAGGTTCTGTCTGTCCACCCAGGCTGGAAACCCAGTGGCAAA	103523	
QY	425	TCATAGTCTCTGTAACCTCAACTCTGGCTCAAGTGATCCTACTACCTCAGCCTCCA	484	
Db	103524	TCTCAGGTCACTACAACCTCTGCTGTGGTTCAGCAATTCCTCGCTCAGCCTCT	103583	
QY	485	GAGTAGCTAGGACTACAGGCACACACAGCCATCTGGCTAAATTTTTTTTAAATTTTC	544	
Db	103584	GAGTAGCTAGGACTACAGGCAGTGTCCACACACACCAATTAATTTTATATTTAATA	103643	
QY	545	ATTTTATGTTATTCATTTCT-----TCTTTTTTTGTTGTTGTTTGTAGATAGG	595	
Db	103644	GAGATGGGGTTTCACTATGTGGCCAGGCTGGCTTTTCTTTGTTGTTTGTAGACAGG	103703	
QY	596	GTCTCATTGTTTACCCAGGCTGGAGGCGATGCGATGGTGCACAGCTGA--GCAGCCTTG	653	
Db	103704	GTCTTACTCTCTTTCCAGGCTAGAGTGCAATCATGGCTCACTGCGAGCTCA	103763	
QY	654	ACTTCCTGGGCTCAAGTGATCTCTCTGCTCAGCTCCCAAGTAGCTGGGACTACAACA	713	
Db	103764	ACCTCTTAGGCTCAAGCAATCTCCACCTCGGTTCCCAAGTGTGGAATTACAGACA	103823	
QY	714	CGTGTCAACATGC---CTGGCTGATATTTTTTTTCTTGAACACAGGGATATCATCTGTTCG	770	
Db	103824	TGAGCTACCAATCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGAGTCTGCTCTGTTCG	103883	
QY	771	CCAGGCTGGAGTACAGTGGCGGTATATATAGTCTACTGAGCCTCCCTCTGGGCTCAAG	830	
Db	103884	CCAGGCTGGATTGCACTGGCAACACCTCGGCTCTACTGCACTCTGCTCTGCTGCTCAAG	103943	
QY	831	CAATCCGCTGCGCTCAGCATCTCAGTAGCTGGGACTACAGGCTTGTGCCACCCAGGCCCA	890	
Db	103944	CAATTCCTGCTCAGCATCTCAGTAGCTGGGACTACAGGCTTGTGCCACCCAGGCCCA	104003	


```
repeat_region 12464..12514 /rpt_family="L1"
repeat_region 12515..12822 /rpt_family="Alu"
repeat_region 12823..12862 /rpt_family="L1"
repeat_region 12857..13096 /rpt_family="L1"
repeat_region 13101..13330 /rpt_family="MaLR"
repeat_region 13331..13634 /rpt_family="Alu"
repeat_region 13635..13692 /rpt_family="MaLR"
repeat_region 13704..13990 /rpt_family="L1"
repeat_region 14082..14387 /rpt_family="Alu"
repeat_region 14402..14704 /rpt_family="Alu"
repeat_region 14797..15095 /rpt_family="Alu"
repeat_region 15162..15196 /rpt_family="Mariner"
repeat_region 15606..15681 /rpt_family="L1"
repeat_region 15691..15827 /rpt_family="MER1_type"
repeat_region 15861..16189 /rpt_family="L2"
repeat_region 16426..16446 /rpt_family="CAAT)n"
repeat_region 16447..16730 /rpt_family="Alu"
repeat_region 17072..17179 /rpt_family="L2"
repeat_region 17212..17504 /rpt_family="Alu"
repeat_region 17869..18139 /rpt_family="Alu"
repeat_region 18168..18464 /rpt_family="Alu"
repeat_region 18606..19155 /rpt_family="MaLR"
repeat_region 19454..19767 /rpt_family="Alu"
repeat_region 19807..19949 /rpt_family="MER1_type"
repeat_region 19950..20247 /rpt_family="Alu"

Query Match 10.7%; Score 423.8; DB 9; Length 190508;
Best Local Similarity 60.1%; Pred No. 8.9e-93;
Matches 969; Conservative 0; Mismatches 547; Indels 97; Gaps 12;

QY 562 TCTTTTCTTTTGTGTTGTTTGGATAGAGTCTCAGTTTGTACCCAGGCTGGAG 621
DB 66758 TCTGATTTTGTGTTTGTGTTTGGACAGGGTTTCACTCTGTCAACCAGGCTGGAG 66699

QY 622 GCGAGTGGCATGGTGACAGCTGA--GCGAGCTTGACTTCTGGGCTCAAGTGATCCTCCT 679
DB 66698 TGTAGTGGTGTGATCATGAGCTCACTGCAACCTTGAACCCCTAGGCTCAAGTCATCCTCCT 66639

QY 680 GCTCAGCTCCCAAGTAGCTGGACTACAAACAGTGTCAACATGCTGGCTGA----- 734
DB 66638 GCTTCAGCTCCCAAGTAGCTGGACTATAGCATGTACCAACATGCCCTGGCTATATTTT 66579

QY 735 -----TATTTTCTTTTGAACAGGGTATACATCTGTGTGCCAGGCTGGAGTACAGTG 788
DB 66578 TAATTTTCTTTTGTGAGACTGAGTCTCGCTGTGTGCCAGGCTGGAGTACAGTG 66519

QY 789 GGTATATATAGCTCACTGACGCTCC--CCTCGTGGCTCAAGCAATCGGCTGGCTCAG 847
DB 66519 GGTGACTCATGCTGTATATCCAGCACTTTGGGAAGCCGAGGTGGTGGATCACTTGAGG 65460
```

```
DB 66518 GCGGATCTCGGCTCACTGCAAGCTCCGCTCCCGGGTTACGGCAATTCCTCTGCTCAA 66459
QY 848 CATCCTGAGTAGCTGGGACTACAGGCTTGTGCCACACAGGCCACGATAGTTTAAAAAAT 907
DB 66458 CCTCCGAGTAGCTGGGACTACAGTCCGCCGCCACACAGCCTGGCTAAATTTTTTTTTT 66399
QY 908 GATTTTGGTATAGAGGAGGTCTTGTATGTCTAGGCTGTATTTTATTTGTGTAGAC 967
DB 66398 TATTTTGTAGAGACAGGGGTTTCACTGTGTAGCAGGATGGTCTCGATCTTCTGACC- 66340
QY 968 AAGTCTCACTATGTTGCCATGATCCGCCACCTCCACTTCCCAAGTGTCACTTATC 1027
DB 66339 -----TTGTGATCCACCTCGCTCGGCTCCCAAGTGTGGGATATACA 66297
QY 1028 TGTTTCATTAGTCAGTTGACAGACATTTTAGTGTGTTTCCACTTTTGTACCAATATGAATA 1087
DB 66296 GCGTGAGCCACTGCACCCAGCCTAATTTTAAATTTTCTGTAGAGACAGGGTTTCGCTA 66237
QY 1088 TACTCCAGTGAATATTCATGTATACATTTGTGTGGGATATGTTTCAATTTCTGTGGGT 1147
DB 66236 TGTTCGCCAGGCTGGTACTGAA-----CTACTGGGCTCAAGCGATCCTCTGTCTGG 66184
QY 1148 TTATATCTAGGAGTGGAAATTCGTGATCCCGGGTAAATATTTTGTACAGCAGATCAGGG 1207
DB 66183 CCACCCAAAGTGTGAGATTACAGCAATGAGCTACCATGCTGGCTG----- 66136
QY 1208 GAAGAAAAAATTTGGAAAAATGAAGCATGTTTAGAAATCAGCAAGAGTGCAGGGTTTTTC 1267
DB 66135 -----AATTTGAAAAATTAATTCAGATTAATTTTAAGTTTATATGTAATTAATAC 66083
QY 1268 GGAGTTTATTTATATTCGTGTGACAAATGTGCAGTTTGTATGAAAGATACAAATATACT 1327
DB 66082 AGCTTAGTTCTCTGAAAAATCAGCAACTGAAGTGTAAATTTGGTATAGGTGGTAACTCAAGT 66023
QY 1328 AAGTGAAGTGAAGTAATTAAGCTGGAAATAGGCGTTTCAGAGTAAATCATGAAGCACTT 1387
DB 66022 AAGAAAAATGTTAGAGACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 65975
QY 1388 TGAATACCAAAATTAAGGAGCTTGGCTGTAACAAAAATAAATAAATAAATAAATAAATAA 1447
DB 65974 TCTGTAATATAACAACATTTTCCATATACAGGGTAAATCTCTGGGTTTTTTTGTGTGT 65915
QY 1448 TTTTGTGAGAAAGAGTCTTGTCTTTTCACTTGTGAGGCTGGAGGAGTGTGTGATCTC 1507
DB 65914 TTTTGTGAGATGGAGTCTTGTCTGTCAACCAGGCTGGAGTGCATGCAATGCAATCTC 65855
QY 1508 AGCTCACTGCAACTTTGCGCTCCCGGTTCAAGCAATTTCTCTGCTTCAGGCTCCCAAGT 1567
DB 65854 GGCTCACTGCAACCTCTGCTCCCAAGTTCAAGTATCTCTCTGCTCAGGCTCCTGGAT 65795
QY 1568 AGCTGGGACTACAGGCACTTCCCAACCATGCGCAAGTGAATTTT-GTATTTTGTAGAGA 1626
DB 65794 AGCTAGGACTACAGGCACTGCGCCACCGCTGGCTCATTTTGTATTTTGTAGAGA 65735
QY 1627 TGGGATTTCACTTTGTGGCCAAAGTGTCTCAAACTTTTGTCTGTCTATAATTTGTGAA 1686
DB 65734 CGGGGTTTACCATGTTAGCCAGGCTGGTCTCAAAAC-TCITGACCTCAAGGATCTGCT 65676
QY 1687 CTATTTGTTCTTTTGTCTGAGGTAGGGCCCCAGACCAAAAAATAAATCTTAGAATCCA 1746
DB 65675 GCCTCTGCTCAGCTTCCCAAGTGTGGGATTTACAGCGGTGACCAACCATGCCAGCA 65616
QY 1747 AATCAGTGTGTGTTGACCACTGTCTGTTGAGAACCAACAGTGTGACAGGCGCTCAGG 1806
DB 65615 AAACCTTATTTAT-----AATCAATCAGATATCTCTAT----- 65579
QY 1807 AGTAGAGTGTCTCTGCTCGAAAGAGAAATAGATGAATATTTCTCCGGGCCAGGCT 1866
DB 65578 -GTGATTTCTCTGTTTACTTTCAAAAAAATTTGTGAAGAAATAAGCTATGCCCCGGGAT 65520
QY 1867 GGTGCTCATCCTGTATATCCAGCACTTTGGGAGGCCAAGGCAATGTGGATCACCTGAGG 1926
DB 65519 GGTGACTCATGCTGTATATCCAGCACTTTGGGAAGCCGAGGTGGTGGATCACTTGAGG 65460
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:18 ; Search time 2121.36 Seconds
(without alignments)
7926.209 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3958

Perfect score: 3958

Sequence: 1 aggaattcaccatttaaat.....gggacctccagcaccgcgg 3958

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_29Jan04.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002s.*

7: geneseqn2003as.*

8: geneseqn2003bs.*

9: geneseqn2003cs.*

10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3958	100.0	4150	ABL57445	ABL57445 Human ste
2	1799.8	45.5	1970	ABK94819	ABK94819 Human ste
3	498.4	12.6	617	AAL26627	AAL26627 Human ste
4	386	9.8	8880	AAK83253	AAK83253 Human imm
5	375.4	9.5	92638	ABQ88096	ABQ88096 Human ost
6	368.8	9.3	5491	AAK69044	AAK69044 Human imm
7	366.4	9.3	9742	AAK42015	AAK42015 Genomic s
8	356.2	9.0	84607	AAK90847	AAK90847 Human PAC
9	349.8	8.8	12394	AAL14749	AAL14749 Human gly
10	349	8.8	13216	AAL05122	AAL05122 Human rep
11	349	8.8	13216	AB198014	AB198014 Human tes
12	348.6	8.8	396	ABV52505	ABV52505 Human pro
13	347.4	8.8	29163	AAL05121	AAL05121 Human rep
14	347.4	8.8	29163	AB198013	AB198013 Human tes
15	347	8.8	6437	AAK42014	AAK42014 Genomic s
16	345	8.7	84495	AAK20588	AAK20588 Human met
17	343.2	8.7	128978	ABK83459	ABK83459 Human cDN
18	343.2	8.7	128978	AAD54587	AAD54587 Human LIM
19	342.8	8.7	26928	AAL232184	AAL232184 Human pro
20	342.8	8.7	26928	ABN95780	ABN95780 Gene #227
21	342.8	8.7	26928	ABQ77403	ABQ77403 Human F2
22	342.8	8.7	26928	ADB88550	ADB88550 Human Fac
23	337.2	8.5	3273	ADB83193	ADB83193 Human cDN

C	24	337	8.5	109906	6	ABK94411	DNA encod
	25	335.2	8.5	128978	6	ABK83459	Human cDN
	26	335.2	8.5	128978	7	AAD54587	Human LIM
C	27	333.8	8.4	183610	7	ACF62736	Cancer ba
	28	333.8	8.4	183610	7	ADB20851	MRP1 base
C	29	333.8	8.4	183610	9	ADB87940	Human UGT
	30	333.8	8.4	183610	9	ADB87940	Human MDR
C	31	333.8	8.4	183610	9	ADB92114	Human MDR
	32	333.6	8.4	35425	9	ADC87616	Human GPC
	33	332	8.4	25715	9	AAS33462	DNA encod
C	34	329.6	8.3	9372	4	AAL36684	Human mus
	35	329.6	8.3	9372	4	AAL36684	Human mus
C	36	329.6	8.3	9372	7	ABX59672	CDNA enco
	37	329.6	8.3	9372	7	ABX59672	CDNA enco
C	38	328.6	8.3	35425	9	ADC87616	Mouse BLM
	39	328.4	8.3	96593	9	ADC85367	Human BLM
	40	328.4	8.3	96594	8	ADA02888	Human BLM
	41	328.4	8.3	96594	9	ADB72626	Human imm
C	42	328	8.3	8308	4	AAK79820	Human imm
	43	328	8.3	8308	4	AAK68091	Human imm
C	44	327.8	8.3	68355	7	ACF62737	Cancer ba
	45	327.8	8.3	68355	7	ADB20852	MRP1 base

ALIGNMENTS

RESULT 1	
ABL57445	
ID	ABL57445 standard; DNA; 4150 BP.
AC	ABL57445;
XX	
DT	22-AUG-2002 (first entry)
XX	
DE	Human stearoyl-CoA desaturase gene 5' region.
XX	
KW	Stearoyl-CoA desaturase; SCD; enzyme; human; promoter; virucide;
KW	dermatological; cytostatic; immunosuppressive; antiallergic;
KW	antiarthritic; antiinflammatory; cardiovascular; antianaemic;
KW	chromosome 10; gene therapy; gene; ds.
XX	
OS	Homo sapiens.
XX	
PH	Key Location/Qualifiers
FT	Misc_feature
FT	1..3958
FT	/tag= a
FT	/note= "region specifically described in Claim 1"
FT	Misc_feature
FT	1..3910
FT	/tag= b
FT	/note= "region specifically described in Claim 2"
FT	protein_bind
FT	complement(105..108)
FT	/tag= c
FT	protein_bind
FT	126..129
FT	/tag= d
FT	/bound_moiety= "HNF-3/Fkh-1"
FT	protein_bind
FT	172..175
FT	/tag= e
FT	/bound_moiety= "SREBP"
FT	protein_bind
FT	complement(255..258)
FT	/tag= f
FT	/bound_moiety= "NF-Y"
FT	protein_bind
FT	complement(263..266)
FT	/tag= g
FT	/bound_moiety= "IRF-2"
FT	protein_bind
FT	complement(461..464)
FT	/tag= h
FT	/bound_moiety= "SREBP"
FT	protein_bind
FT	complement(541..544)
FT	/tag= i
FT	/bound_moiety= "IRF-1"
FT	protein_bind
FT	complement(669..672)

FT	protein_bind	/tag= j	bound moiety= "SREBP"	FT	protein_bind	bound moiety= "RAR-alpha"
FT	protein_bind	991. .994		FT	protein_bind	2359. .2362
FT	protein_bind	/tag= k		FT	protein_bind	/tag= ai
FT	protein_bind	bound moiety= "SREBP"		FT	protein_bind	bound moiety= "NF-1"
FT	protein_bind	complement (1037. .1040)		FT	protein_bind	complement (2365. .2368)
FT	protein_bind	/tag= l		FT	protein_bind	/tag= aj
FT	protein_bind	bound moiety= "AP-1"		FT	protein_bind	bound moiety= "ER"
FT	protein_bind	complement (1073. .1076)		FT	protein_bind	complement (2382. .2385)
FT	protein_bind	/tag= m		FT	protein_bind	/tag= ak
FT	protein_bind	bound moiety= "RAR-alpha"		FT	protein_bind	bound moiety= "NF-1"
FT	protein_bind	1240. .1243		FT	protein_bind	complement (2414. .2417)
FT	protein_bind	/tag= n		FT	protein_bind	/tag= al
FT	protein_bind	bound moiety= "C/EBP-beta"		FT	protein_bind	bound moiety= "Rel"
FT	protein_bind	1410. .1413		FT	protein_bind	2445. .2450
FT	protein_bind	/tag= o		FT	protein_bind	/tag= am
FT	protein_bind	bound moiety= "NF-1"		FT	protein_bind	bound moiety= "AP-1"
FT	protein_bind	complement (1423. .1426)		FT	protein_bind	complement (2452. .2455)
FT	protein_bind	/tag= p		FT	protein_bind	/tag= an
FT	protein_bind	bound moiety= "HNF3-beta"		FT	protein_bind	bound moiety= "NF-1"
FT	protein_bind	complement (1500. .1503)		FT	protein_bind	complement (2533. .2536)
FT	protein_bind	/tag= q		FT	protein_bind	/tag= ao
FT	protein_bind	bound moiety= "SREBP"		FT	protein_bind	bound moiety= "GR"
FT	protein_bind	complement (1587. .1590)		FT	protein_bind	complement (2726. .2729)
FT	protein_bind	/tag= r		FT	protein_bind	/tag= ap
FT	protein_bind	bound moiety= "NF-kB"		FT	protein_bind	bound moiety= "CREB"
FT	protein_bind	1628. .1631		FT	protein_bind	complement (2744. .2747)
FT	protein_bind	/tag= s		FT	protein_bind	/tag= aq
FT	protein_bind	bound moiety= "NF-kB"		FT	protein_bind	bound moiety= "ER"
FT	protein_bind	complement (1632. .1634)		FT	protein_bind	2862. .2865
FT	protein_bind	/tag= t		FT	protein_bind	/tag= ar
FT	protein_bind	bound moiety= "IRF-1"		FT	protein_bind	bound moiety= "NF-1"
FT	protein_bind	1713. .1716		FT	protein_bind	complement (2917. .2920)
FT	protein_bind	/tag= u		FT	protein_bind	/tag= as
FT	protein_bind	bound moiety= "AP-2"		FT	protein_bind	bound moiety= "PPAR-alpha"
FT	protein_bind	complement (1729. .1732)		FT	protein_bind	complement (2942. .2945)
FT	protein_bind	/tag= v		FT	protein_bind	/tag= at
FT	protein_bind	bound moiety= "HNF3-beta"		FT	protein_bind	bound moiety= "Oct-1"
FT	protein_bind	1833. .1836		FT	protein_bind	2958. .2961
FT	protein_bind	/tag= w		FT	protein_bind	/tag= au
FT	protein_bind	bound moiety= "IRF-2"		FT	protein_bind	bound moiety= "CREB/ATF"
FT	protein_bind	complement (1863. .1866)		FT	protein_bind	complement (2965. .2968)
FT	protein_bind	/tag= x		FT	protein_bind	/tag= av
FT	protein_bind	bound moiety= "Whn"		FT	protein_bind	bound moiety= "STAT"
FT	protein_bind	1917. .1920		FT	protein_bind	complement (2979. .2982)
FT	protein_bind	/tag= y		FT	protein_bind	/tag= aw
FT	protein_bind	bound moiety= "SREBP"		FT	protein_bind	bound moiety= "C/EBP-beta"
FT	protein_bind	1925. .1928		FT	protein_bind	3056. .3059
FT	protein_bind	/tag= z		FT	protein_bind	/tag= ax
FT	protein_bind	bound moiety= "T3R, RAR-alpha"		FT	protein_bind	bound moiety= "NF-kB"
FT	protein_bind	complement (2087. .2090)		FT	protein_bind	complement (3062. .3076)
FT	protein_bind	/tag= aa		FT	protein_bind	/tag= ay
FT	protein_bind	bound moiety= "Pax-6"		FT	protein_bind	bound moiety= "NF-kB"
FT	protein_bind	2093. .2096		FT	protein_bind	complement (3065. .3068)
FT	protein_bind	/tag= ab		FT	protein_bind	/tag= az
FT	protein_bind	bound moiety= "SREBP"		FT	protein_bind	bound moiety= "SP-1"
FT	protein_bind	complement (2116. .2119)		FT	protein_bind	3131. .3134
FT	protein_bind	/tag= ac		FT	protein_bind	/tag= ba
FT	protein_bind	bound moiety= "AP-2"		FT	protein_bind	bound moiety= "Ets-1"
FT	protein_bind	complement (2179. .2182)		FT	protein_bind	complement (3161. .3164)
FT	protein_bind	/tag= ad		FT	protein_bind	/tag= bb
FT	protein_bind	bound moiety= "CHOP"		FT	protein_bind	bound moiety= "Whn"
FT	protein_bind	2199. .2202		FT	protein_bind	complement (3162. .3166)
FT	protein_bind	/tag= ae		FT	protein_bind	/tag= bc
FT	protein_bind	bound moiety= "RAR-alpha"		FT	protein_bind	bound moiety= "Myc/Max"
FT	protein_bind	complement (2231. .2234)		FT	protein_bind	3173. .3176
FT	protein_bind	/tag= af		FT	protein_bind	/tag= bd
FT	protein_bind	bound moiety= "NF-1"		FT	protein_bind	bound moiety= "SP-1"
FT	protein_bind	2245. .2248		FT	protein_bind	32

```
FT protein_bind complement (3245..3248)
FT /*tag= bg
FT /bound_moiety= "Whn"
FT protein_bind complement (3276..3279)
FT /*tag= bh
FT /bound_moiety= "NF1"
FT protein_bind 3291..3294
FT /*tag= bi
FT /bound_moiety= "Myc/Max"
FT protein_bind 3292..3295

Query Match      100.0%; Score 3958; DB 6; Length 4150;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGAATTCATCATTAATTAATCATACAAATTAATAGGCTTTAGTATATATCAAGGTTGTC 60
Db 1 AGGAATTCATCATTAATTAATCATACAAATTAATAGGCTTTAGTATATATCAAGGTTGTC 60

Qy 61 ATCCATGACATCCATTTAGACAGTTTATTAATAGGCTTTAGTATATATCAAGGTTGTC 120
Db 61 ATCCATGACATCCATTTAGACAGTTTATTAATAGGCTTTAGTATATATCAAGGTTGTC 120

Qy 121 AGCCATACCCCCCAACATCCTCCATCCTCTTCCAAAGCCCTGGGCAACCAACCAATCTAC 180
Db 121 AGCCATACCCCCCAACATCCTCCATCCTCTTCCAAAGCCCTGGGCAACCAACCAATCTAC 180

Qy 181 TTTCTGCTCTATATAATTTGCCAATTTCTGACATTTTCATATAAATGGAAAGCAACAACAT 240
Db 181 TTTCTGCTCTATATAATTTGCCAATTTCTGACATTTTCATATAAATGGAAAGCAACAACAT 240

Qy 241 GTGAGACTTTTGACTGCTGCTTTCACCTAGCATTTCTATTTTAAAGGCTCAATATGTTA 300
Db 241 GTGAGACTTTTGACTGCTGCTTTCACCTAGCATTTCTATTTTAAAGGCTCAATATGTTA 300

Qy 301 CAGTACTTAGCAGTACTTCACTTTTATTTCTCAAAATGGTATTTCCACTGTGGGTAT 360
Db 301 CAGTACTTAGCAGTACTTCACTTTTATTTCTCAAAATGGTATTTCCACTGTGGGTAT 360

Qy 361 CCCATATCATATTTAGAGACAGGTTCTCATCTCTGTCAACAGGCTGGAGTGCAGTGGC 420
Db 361 CCCATATCATATTTAGAGACAGGTTCTCATCTCTGTCAACAGGCTGGAGTGCAGTGGC 420

Qy 421 ACAATCATAGTCACTGTAACTCAAACTCTGGCTCAAGTGATCTTACCTCAGCC 480
Db 421 ACAATCATAGTCACTGTAACTCAAACTCTGGCTCAAGTGATCTTACCTCAGCC 480

Qy 481 TCCAGAGTAGCTAGGACTACAGGCACACACAGCCATACCTGGCTAAATTTTTTTTAAAT 540
Db 481 TCCAGAGTAGCTAGGACTACAGGCACACACAGCCATACCTGGCTAAATTTTTTTTAAAT 540

Qy 541 TTTCAATTTATGATATCAATTTCTTTCTTTTGTGTTGTTGTTGTTGATAGGCTCTC 600
Db 541 TTTCAATTTATGATATCAATTTCTTTCTTTTGTGTTGTTGTTGTTGATAGGCTCTC 600

Qy 601 ACTTTGTTACCCAGGCTGAGGAGTGGCATGTGACAGCTGAGCAGCTTGACTTCCT 660
Db 601 ACTTTGTTACCCAGGCTGAGGAGTGGCATGTGACAGCTGAGCAGCTTGACTTCCT 660

Qy 661 GGGCTCAAGTGATCCTCCTGCTCAGCCCTCCCAAGTAGCTGGGACTACAAACACGTTGCA 720
Db 661 GGGCTCAAGTGATCCTCCTGCTCAGCCCTCCCAAGTAGCTGGGACTACAAACACGTTGCA 720

Qy 721 CCATGCTGGCTGATATTTTCTTGAACAGGGTATCACTCTGTTGCCAGGCTGGA 780
Db 721 CCATGCTGGCTGATATTTTCTTGAACAGGGTATCACTCTGTTGCCAGGCTGGA 780

Qy 781 GTACAGTGGCTGATTAATAGCTCACTGAGCCTCCCTCCCTGGGCTCAAGCAATCCGCTG 840
Db 781 GTACAGTGGCTGATTAATAGCTCACTGAGCCTCCCTCCCTGGGCTCAAGCAATCCGCTG 840

Qy 841 GCCTCAGCATCCTGAGTAGTGGGACTACAGGCTTTGTGCCACAGGCCAGCTAAGTTTTT 900
Db 841 GCCTCAGCATCCTGAGTAGTGGGACTACAGGCTTTGTGCCACAGGCCAGCTAAGTTTTT 900
```

```
Db 841 GCCTCAGCATCCTGAGTAGTGGGACTACAGGCTTTGTGCCACAGGCCAGCTAAGTTTTT 900
Qy 901 AAAAAATGATTTTTTGGTATAGAGAGGCTTCTGCTATGTTGCTCAGGCTGATTTTTATG 960
Db 901 AAAAAATGATTTTTTGGTATAGAGAGGCTTCTGCTATGTTGCTCAGGCTGATTTTTATG 960
Qy 961 TTGAGACAAGGCTCTCACTATGTTGCCATGATCCCCCACTCCCACTTCCCAAAGTGCTCA 1020
Db 961 TTGAGACAAGGCTCTCACTATGTTGCCATGATCCCCCACTCCCACTTCCCAAAGTGCTCA 1020
Qy 1021 TCTTATCTGTTCAATGCTCAGTTCACAGACATTTAGTGTGTTTCCATTTTGTGACCATTA 1080
Db 1021 TCTTATCTGTTCAATGCTCAGTTCACAGACATTTAGTGTGTTTCCATTTTGTGACCATTA 1080
Qy 1081 TGAATAATCTCCAGTGAATTTTATGATATACATTTGTTGGGCATATGTTTCAATTTCT 1140
Db 1081 TGAATAATCTCCAGTGAATTTTATGATATACATTTGTTGGGCATATGTTTCAATTTCT 1140
Qy 1141 GTTGGGTTTATATCTAGGAGTGAATTTGCTGGATCCCGGTTAAATTTTGACAGGCAGAG 1200
Db 1141 GTTGGGTTTATATCTAGGAGTGAATTTGCTGGATCCCGGTTAAATTTTGACAGGCAGAG 1200
Qy 1201 TTCAGGGGAAGAAAACTTGGGAAATGAAGCATTTTGTAGAAATCAGCAAGTGCAGGG 1260
Db 1201 TTCAGGGGAAGAAAACTTGGGAAATGAAGCATTTTGTAGAAATCAGCAAGTGCAGGG 1260
Qy 1261 GTTTTTCGGAGTTTATTTTATTTCTGTTGACAAAATGTCAGTTTGTAGAGATACAG 1320
Db 1261 GTTTTTCGGAGTTTATTTTATTTCTGTTGACAAAATGTCAGTTTGTAGAGATACAG 1320
Qy 1321 TTATACTAAGTGAGAGTGAGAAATTAAGGCTGGAATAGGGGTTTCCAGAGTAAATCATGA 1380
Db 1321 TTATACTAAGTGAGAGTGAGAAATTAAGGCTGGAATAGGGGTTTCCAGAGTAAATCATGA 1380
Qy 1381 AGCACTTTGAAATCAAAATTAAGAGCTTGGCTGTAAACAAAATAATAAAAAATCAAA 1440
Db 1381 AGCACTTTGAAATCAAAATTAAGAGCTTGGCTGTAAACAAAATAATAAAAAATCAAA 1440
Qy 1441 TTTTTCCTGAGTCACTGCAACTTTTCGCTCCCGGTTTCAAGCAATTTCTCCTGCTCAGCCT 1500
Db 1441 TTTTTCCTGAGTCACTGCAACTTTTCGCTCCCGGTTTCAAGCAATTTCTCCTGCTCAGCCT 1500
Qy 1501 TGATCTCAGTCACTGCAACTTTTCGCTCCCGGTTTCAAGCAATTTCTCCTGCTCAGCCT 1560
Db 1501 TGATCTCAGTCACTGCAACTTTTCGCTCCCGGTTTCAAGCAATTTCTCCTGCTCAGCCT 1560
Qy 1561 CCCAAGTAGCTGGGACTACAGGCACCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Db 1561 CCCAAGTAGCTGGGACTACAGGCACCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Qy 1621 TAGAGATGGGATTTTCACTTTGTTGGCCAGCTGGTCTCAAACTTTTGTCTGCTGCTGCTGCT 1680
Db 1621 TAGAGATGGGATTTTCACTTTGTTGGCCAGCTGGTCTCAAACTTTTGTCTGCTGCTGCTGCT 1680
Qy 1681 TTGTAACATTTGTTTCTTTTCTGCTGAGGTAGGGCCCCCAGACCAAAAAATAAATCTTAG 1740
Db 1681 TTGTAACATTTGTTTCTTTTCTGCTGAGGTAGGGCCCCCAGACCAAAAAATAAATCTTAG 1740
Qy 1741 AATCCAAATCAGTGTGTTGTTGACCACTGCTCAGTTCAGAGAACCAAGTGTGACAGGGC 1800
Db 1741 AATCCAAATCAGTGTGTTGTTGACCACTGCTCAGTTCAGAGAACCAAGTGTGACAGGGC 1800
Qy 1801 CTCAGGAGTAGGATGATCTCTGCTCGAAAGAGAAATAGAAATGAAATATTTCTCGGGGCC 1860
Db 1801 CTCAGGAGTAGGATGATCTCTGCTCGAAAGAGAAATAGAAATGAAATATTTCTCGGGGCC 1860
Qy 1861 AGGCGTGGTGGCTCATGCTCTAATCCAGCACTTTGGAGGCCAAGCAATGTGATCAC 1920
Db 1861 AGGCGTGGTGGCTCATGCTCTAATCCAGCACTTTGGAGGCCAAGCAATGTGATCAC 1920
Qy 1921 CTGAGGTGAGAGTTCAAAACAGGCTGGCCAAATGTTGAAACCCCTCTCTACTAAA 1980
Db 1921 CTGAGGTGAGAGTTCAAAACAGGCTGGCCAAATGTTGAAACCCCTCTCTACTAAA 1980
```

QY 1981 ATACAAAAATTTAGCTAAGTGTGGTGGCCATGCCTGTAATCCAGCTACTTTGGAGGGT 2040
 Db 1981 ATACAAAAATTTAGCTAAGTGTGGTGGCCATGCCTGTAATCCAGCTACTTTGGAGGGT 2040
 QY 2041 GAGGACAGAGAAATTTCTTGAACCCCGGAGGACAGAGTTCAGTGAAGCGAGATCAACCA 2100
 Db 2041 GAGGACAGAGAAATTTCTTGAACCCCGGAGGACAGAGTTCAGTGAAGCGAGATCAACCA 2100
 QY 2101 CTGCACTCCAGCTGGGGAGAGAGAGACTTCCTCTCAAAAAAACAACAAAAAAGA 2160
 Db 2101 CTGCACTCCAGCTGGGGAGAGAGAGACTTCCTCTCAAAAAAACAACAAAAAAGA 2160
 QY 2161 ATTAAGCAATTTAGCAATTTGACAGAGAACTTGAAGGGGTGAGACCACTGACATTC 2220
 Db 2161 ATTAAGCAATTTAGCAATTTGACAGAGAACTTGAAGGGGTGAGACCACTGACATTC 2220
 QY 2221 TGTGCCACATGCAAGTACTTCTGAGGCATGACTGGATGAGTGTCCACATCTGAATCA 2280
 Db 2221 TGTGCCACATGCAAGTACTTCTGAGGCATGACTGGATGAGTGTCCACATCTGAATCA 2280
 QY 2281 TCCAGTCTTGTTCAGAACTTTTCAACCGGACAGGGAGCCAGACTGGAATGACATCTCCT 2340
 Db 2281 TCCAGTCTTGTTCAGAACTTTTCAACCGGACAGGGAGCCAGACTGGAATGACATCTCCT 2340
 QY 2341 GGTCACTGGCCAGAGAGTGGGCTTGAACCTGAGACAGTGGCCACAAAGGAGTGTCTT 2400
 Db 2341 GGTCACTGGCCAGAGAGTGGGCTTGAACCTGAGACAGTGGCCACAAAGGAGTGTCTT 2400
 QY 2401 AGTCTACCTCCAGGAAATCCAGAGTGTGTCTTCTGGGAAGTGAATCATTTGGCGAG 2460
 Db 2401 AGTCTACCTCCAGGAAATCCAGAGTGTGTCTTCTGGGAAGTGAATCATTTGGCGAG 2460
 QY 2461 CACTCGGATTTTCTCTCTTCCAGGGAAGATCCTTAGGGCAGTATTTGGGAAGACA 2520
 Db 2461 CACTCGGATTTTCTCTCTTCCAGGGAAGATCCTTAGGGCAGTATTTGGGAAGACA 2520
 QY 2521 TGGGATGAAGGACACCGGTGAATGCATAGCTGCTGCTGCTCTCATGGTAA 2580
 Db 2521 TGGGATGAAGGACACCGGTGAATGCATAGCTGCTGCTGCTCTCATGGTAA 2580
 QY 2581 GGCTCCTCAGACACGGAAGATGGGGCACAGGACAGATCAGTAGGGTCAAGCATC 2640
 Db 2581 GGCTCCTCAGACACGGAAGATGGGGCACAGGACAGATCAGTAGGGTCAAGCATC 2640
 QY 2641 TCAGGACCGAGGCAATATGCTCTGAGCAGGGATTAAGAGCTTGGGCTCTCATATGGT 2700
 Db 2641 TCAGGACCGAGGCAATATGCTCTGAGCAGGGATTAAGAGCTTGGGCTCTCATATGGT 2700
 QY 2701 GTTCTGGGCTCACTGCCAGCTCCGCTCACTTACTGTTGCTGTGACCATGGGCAAGTTA 2760
 Db 2701 GTTCTGGGCTCACTGCCAGCTCCGCTCACTTACTGTTGCTGTGACCATGGGCAAGTTA 2760
 QY 2761 TTCCATCTCTCCATATCTCTTCTCCTCACTTTTAAATGGAATTAATGGGTAACCACTCC 2820
 Db 2761 TTCCATCTCTCCATATCTCTTCTCCTCACTTTTAAATGGAATTAATGGGTAACCACTCC 2820
 QY 2821 CAGGGTCAAGAGAGCTTACAGAAACGATCTTCTGTAATGGCTGCAATTAATTC 2880
 Db 2821 CAGGGTCAAGAGAGCTTACAGAAACGATCTTCTGTAATGGCTGCAATTAATTC 2880
 QY 2881 AATACCTGCCAGCTATTTCTTATTCACATCCAGCCCTTTCCGCTGCTGCTGGGTGAAA 2940
 Db 2881 AATACCTGCCAGCTATTTCTTATTCACATCCAGCCCTTTCCGCTGCTGCTGGGTGAAA 2940
 QY 2941 CACATGTCAGTGTTCCTGACGGTTTCCACAAAGAGATTCACAAATTAACCTGCCAG 3000
 Db 2941 CACATGTCAGTGTTCCTGACGGTTTCCACAAAGAGATTCACAAATTAACCTGCCAG 3000
 QY 3001 TCTGAAGATCTCCAAAACATCCCGACGATCCTGGAGCGGGCTTGGGATGGAC 3060
 Db 3001 TCTGAAGATCTCCAAAACATCCCGACGATCCTGGAGCGGGCTTGGGATGGAC 3060

QY 3061 TGCCCGCCGGGTCTCTGAACAGGATGCTGCGCGCAGGCACACACACACAGCCAGCCTG 3120
 Db 3061 TGCCCGCCGGGTCTCTGAACAGGATGCTGCGCGCAGGCACACACACACAGCCAGCCTG 3120
 QY 3121 TGTGTGCGCGCGGAGTCCGGTCCGGTGGAGCAGCGCTGCTGCTGGTGGCGGGC 3180
 Db 3121 TGTGTGCGCGCGGAGTCCGGTCCGGTGGAGCAGCGCTGCTGCTGGTGGCGGGC 3180
 QY 3181 AGAGCCATTTTTCGACGGCTACCGAGCCCGCGCTGCGCCGGGAGGAGCGGGCT 3240
 Db 3181 AGAGCCATTTTTCGACGGCTACCGAGCCCGCGCTGCGCCGGGAGGAGCGGGCT 3240
 QY 3241 TCCCGCTCCCAAGCTCCAGATCCTGGGTGGTCCACAGCTCCTGCTCCACGCGCTG 3300
 Db 3241 TCCCGCTCCCAAGCTCCAGATCCTGGGTGGTCCACAGCTCCTGCTCCACGCGCTG 3300
 QY 3301 GGGGACCGGGAAGACGGGACGGGATGTTAGTGTGGGCGCCCGCGAGGGTTCACACT 3360
 Db 3301 GGGGACCGGGAAGACGGGACGGGATGTTAGTGTGGGCGCCCGCGAGGGTTCACACT 3360
 QY 3361 GTTCTCTGAGAACTTCCAGTGTCCACCGCTTCTCGTGTGCGCGGGCGCT 3420
 Db 3361 GTTCTCTGAGAACTTCCAGTGTCCACCGCTTCTCGTGTGCGCGGGCGCT 3420
 QY 3421 CTTGGCTAGGCTCCGCGCCCGCAGCCCAACCGGGTCCCGAGCCCTTCCAGAGAGAA 3480
 Db 3421 CTTGGCTAGGCTCCGCGCCCGCAGCCCAACCGGGTCCCGAGCCCTTCCAGAGAGAA 3480
 QY 3481 GTCCTCGACCGGGATGCGGGCAGAGCCAGCGCGGGTGGAAAGAGTGTGAAGG 3540
 Db 3481 GTCCTCGACCGGGATGCGGGCAGAGCCAGCGCGGGTGGAAAGAGTGTGAAGG 3540
 QY 3541 AGAAACAGAGGGGAGGGGAGCGAGCTGCGCGCAGAGGAAACAGCAGATTGCGCGA 3600
 Db 3541 AGAAACAGAGGGGAGGGGAGCGAGCTGCGCGCAGAGGAAACAGCAGATTGCGCGA 3600
 QY 3601 GCCAATGGCAACCGGACGAGGTGGCAACCAATTTCCCTTCGGCCAATGACGAGCGGA 3660
 Db 3601 GCCAATGGCAACCGGACGAGGTGGCAACCAATTTCCCTTCGGCCAATGACGAGCGGA 3660
 QY 3661 GTTTACAGAGCTCATTAAGCATTTTCCCGAGAGGAGGGGAGGGGAGGGCGGGTGG 3720
 Db 3661 GTTTACAGAGCTCATTAAGCATTTTCCCGAGAGGAGGGGAGGGGAGGGCGGGTGG 3720
 QY 3721 TGTGTGTGCTGCTGCGCAGCATCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3780
 Db 3721 TGTGTGTGCTGCTGCGCAGCATCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3780
 QY 3781 TCTGTCT 3840
 Db 3781 TCTGTCT 3840
 QY 3841 TCGCATTTTCCCGCTTGGCAGCGGATTAAGGGGCTGAGGAAATACCGGACACGGT 3900
 Db 3841 TCGCATTTTCCCGCTTGGCAGCGGATTAAGGGGCTGAGGAAATACCGGACACGGT 3900
 QY 3901 CACCGTGTGCGAGCTTAGCTTTAAATTTCCCGCTTGGGAGCTTCCAGCAGCGGG 3958
 Db 3901 CACCGTGTGCGAGCTTAGCTTTAAATTTCCCGCTTGGGAGCTTCCAGCAGCGGG 3958

RESULT 2
 ABK94819

ID ABK94819 standard; DNA; 1970 BP.

XX ABK94819;

AC 29-AUG-2002 (first entry)

DE Human stearyl-CoA desaturase (SCD) control region.

XX Fatty acid regulated gene; polyunsaturated fatty acid disorder;
 KW PUFA disorder; eczema; cardiovascular disorder; hypertriglyceridaemia;

KW dyslipidaemia; atherosclerosis; coronary artery disease;
KW cerebrovascular disease; peripheral vascular disease; inflammation;
KW sinusitis; asthma; pancreatitis; osteoarthritis; rheumatoid arthritis;
KW acne; body weight disorder; obesity; cachexia; anorexia;
KW psychiatric disorder; cancer; cystic fibrosis; pre-menstrual syndrome;
KW diabetes; diabetic complication; genetic polymorphism; ds.
XX Homo sapiens.
XX W0200240666-82.
XX PD 23-MAY-2002.
XX PF 19-NOV-2001; 2001WO-CA001632.
XX PR 17-NOV-2000; 2000US-0248589P.
XX PA (XENO-) XENON GENETICS INC.
XX PI Winther MD, Goldberg YP, Knickle LC, Haardt M, Allen SJ;
XX PI Ponton A, De Antueno RJ, Jenkins DK, Nwaka SO;
XX DR WPI; 2002-508327/54.
XX Novel isolated polypeptide segment encoded by fat regulated genes, useful
XX PT for diagnosing the presence of or a predisposition for a disorder
XX PT involving fatty acid regulated genes in a subject.
XX PS Claim 23; Fig 26; 225pp; English.
XX The invention describes an isolated polypeptide segment (I) whose genes
XX are fat regulated. (I) or the polynucleotide encoding it (II) are useful
XX for diagnosing the presence of or a predisposition for a disorder
XX involving fatty acid regulated genes in a subject. A composition
XX containing (I) or (II) is useful for treating a disorder involving fatty
XX acid regulated genes, where the disorder is selected from a
XX polyunsaturated fatty acid (PUFA) disorder, eczema, cardiovascular
XX disorders (such as hypertriglyceridaemia, dyslipidaemia, atherosclerosis,
XX coronary artery disease, cerebrovascular disease or peripheral vascular
XX disease), inflammation (such as sinusitis, asthma, pancreatitis,
XX osteoarthritis, rheumatoid arthritis or acne), body weight disorders
XX (such as obesity, cachexia or anorexia), psychiatric disorders, cancer,
XX cystic fibrosis, pre-menstrual syndrome, diabetes, and diabetic
XX complications. (I) or (II) is useful as research agent and materials for
XX discovery of treatments and diagnostics for a disease, particularly human
XX disease. (II) is useful for constructing nucleotide probes and primers,
XX for detecting genetic polymorphism, for detecting changes in the level of
XX expression of (II), and as a diagnostic tool. This sequence represents a
XX control region from a gene encoding a fatty acid regulated protein
XX SQ Sequence 1970 BP; 428 A; 590 C; 588 G; 364 T; 0 U; 0 Other;
Query Match 45.5%; Score 1799.8; DB 6; Length 1970;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1834; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY 2121 CAGAGCGAGACTTCTCTCAAAAAAACAAGAAATTAAGCAAAATAGACATTG 2180
DB 1 GAGAGCGAGACTTCTCTCAAAAAAACAAGAAATTAAGCAAAATAGACATTG 60
QY 2181 CAGAGAGAACTGAAGGGGGTCAGACCAAGTACAGATTCTGTGCCATGCCAAGTACT 2240
DB 61 CAGAGAGAACTGAAGGGGGTCAGACCAAGTACAGATTCTGTGCCATGCCAAGTACT 120
QY 2241 TCTGAGCATGACTGAGTACTGCTCCACATCTGAAATCATCCAGTCTGTTCAGAACTT 2300
DB 121 TCTGAGCATGACTGAGTACTGCTCCACATCTGAAATCATCCAGTCTGTTCAGAACTT 180
QY 2301 TCACACCGGACAGGGAGGAGCAGGACTGGAATGAGTCTCCTGGTCACTGGCCAGAGAGTTG 2360
DB 181 TCACACCGGACAGGGAGGAGCAGGACTGGAATGAGTCTCCTGGTCACTGGCCAGAGAGTTG 240
QY 2361 GCCTTGACCTGAGACCAAGTGGCCAAACAAGAGAGTGTCTTACCTCCACGGAATC 2420

DB 241 GCC-TGACCCCTGAGACAGTGGCCCAACAAGAGAGTCTCTTAGTCTACCTCCAGGAATC 299
QY 2421 CCAGGTGTTGCTTCTCTGGGAAGTGAATCATTTGGCGCAGCACTCCGTAATTTTCTCTCT 2480
DB 300 CCAGGTGTTGCTTCTCTGGGAAGTGAATCATTTGGCGCAGCACTCCGTAATTTTCTCTCT 359
QY 2481 TCCAGGGGAGGATCCTTAGGGCAGTATTTGGGAAAGACATGGGCATCGAAGGACACCGG 2540
DB 360 TCCAGGGGAGGATCCTTAGGGCAGTATTTGGGAAAGACATGGGCATCGAAGGACACCGG 419
QY 2541 GTGAATGCATAGCCTGCTTCTGAGCTCTCATGTGTAAGGCTCTTACACACACGGA 2600
DB 420 GTGAATGCATAGCCTGCTTCTGAGCTCTCATGTGTAAGGCTCTTACACACACGGA 479
QY 2601 AGATGGGGCACAAGGACAGATCAGTAGGGTCAGAGCATCTCAGGACCGGAGGCAATAT 2660
DB 480 AGATGGGGCACAAGGACAGATCAGTAGGGTCAGAGCATCTCAGGACCGGAGGCAATAT 539
QY 2661 GGTCTCAGCAGGAGTAAAGAGCTTGGGCTCTCATATGGTGTCTTCTGGGCTCAATGCGCA 2720
DB 540 GGTCTCAGCAGGAGTAAAGAGCTTGGGCTCTCATATGGTGTCTTCTGGGCTCAATGCGCA 599
QY 2721 GCTCCGTCATCTACTGTTGCTGTGACCATGGCAAGTATTCATCTCTCCATATCTCT 2780
DB 600 GCTCCATCCTACTGTTGCTGTGACCATGGCAAGTATTCATCTCTCCATATCTCT 659
QY 2781 TTCTCTCACTTTTAAATGGAAATATGGGGTACCACCTCCAGGGTCACAGAGGGCTTA 2840
DB 660 TTCTCTCACTTTTAAATGGAAATATGGGGTACCACCTCCAGGGTCACAGAGGGCTTA 719
QY 2841 CAGAAAAAGATTTCTGTGAATTTGGCTTGCAGTAATAATCAATACCTGCCAGCTATTCTT 2900
DB 720 CAGAAAAAGATTTCTGTGAATTTGGCTTGCAGTAATAATCAATACCTGCCAGCTATTCTT 779
QY 2901 ATTCCATCATCAAGCCCTTCTGCTGTGTGGTGAAGACATGTGAGTGTTCCTGA 2960
DB 780 ATTCCATCATCAAGCCCTTCTGCTGTGTGGTGAAGACATGTGAGTGTTCCTGA 839
QY 2961 CGGTTCACAAAGAAATTTCCAAATATCAACCTGCCAGTCTGAAGAACTCCCAAAACA 3020
DB 840 CGGTTCACAAAGAAATTTCCAAATATCAACCTGCCAGTCTGAAGAACTCCCAAAACA 899
QY 3021 TCCCGCAGCATCTCTGAGGCGGGCTTGGGATGGAGTCTGCCCGCCGGCTCTGAAC 3080
DB 900 TCCCGCAGCATCTCTGAGGCGGGCTTGGGATGGAGTCTGCCCGCCGGCTCTGAAC 959
QY 3081 AGATGTGCGGCGCAGGACACACACACACAGCCTGTGTGTGGCGCGGAGTCCGG 3140
DB 960 AGATGTGCGGCGCAGGACACACACACAGCCTGTGTGTGGCGCGGAGTCCGG 1019
QY 3141 TGCGGTCCCGGTGAGCAGCGGCTGTGGTGGGGGCGAGAGCCATTGTTCAGGCG 3200
DB 1020 TGCGGTCCCGGTGAGCAGCGGCTGTGGTGGGGGCGAGAGCCATTGTTCAGGCG 1079
QY 3201 TACGAGCCCCCGCGCTCGCCGCGGAGGCGGGCTTCCCGCT-CCCCAAGTCTC 3259
DB 1080 TACGAGTCCCCCGCGCTCGCCGCGGAGGCGGGCTTCCCGCTCCCCAAGTCTC 1139
QY 3260 AGATCTCGGGTGTGCTGCCAGTCTCCCTGCCAGCGCCCTGGGGGAGCGGAGACGGGA 3319
DB 1140 AGATCTCGGGTGTGCTGCCAGTCTCCCTGCCAGCGCCCTGGGGGAGCGGAGACGGGA 1198
QY 3320 CGGAGATGTTAGTGTGGGCGCCCCCGAGGGTTTACACATGTTTCTTGAGAAATTTCC 3379
DB 1199 CGGAGATGTTAGTGTGGGCGCCCCCGAGGGTTTACACATGTTTCTTGAGAAATTTCC 1258
QY 3380 CAGTGTCCACCCCGCTTCTCGGTGTGCCGAGGGCGGGTCTCTGGGTAGGCTCCGGCG 3439
DB 1259 CAGTGTCCACCCCGCTTCTCGGTGTGCCGAGGGCGGGTCTCTGGGTAGGCTCCGGCG 1318
QY 3440 CCAGGCCCAAAACCGGGTCCCCAGCCCTTCCAGAGAGAGTCTCCCGACCGGGATGCC 3499

PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.
 XX
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 PT
 XX
 PS Disclosure; SEQ ID NO 38065; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic

CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 8880 BP; 2388 A; 1807 C; 1844 G; 2841 T; 0 U; 0 Other;
 Query Match 9.8%; Score 386; DB 4; Length 8880;
 Best Local Similarity 54.7%; Pred. No. 2.1e-69;
 Matches 1194; Conservative 0; Mismatches 850; Indels 138; Gaps 16;
 QY 14 TTTAAATCATACAATTTAATGCTTTTGTATATATTCACAG-GTTGTGCAATCAT---CAC 69
 DB 6137 TTTAAAGTGTGCAATTTGAATAATTTTGTACATTTCTAGAGTTGTGCAACCAATTACCA 6196
 QY 70 AATCCATTTTGAACAGTTTTTATCTCCAAATAAACCCTGCAATTCCTTATGCCATCAC 129
 DB 6197 CATCAAGTTTAGACATTTTTTATACCCCAAGAAACCCCTGTAGCTTTAGCTGTAC 6256
 QY 130 CCCCCA--ACATCCTCCATCCTCCTTCCAAAGCCCTGGGCAACCAACCAATCTACTTTCTGT 187
 DB 6257 CTTCCAGTACCACTTCTCTCTCCCTTTCCCTAGGCTATCACTAATAAACTTTCTGT 6316
 QY 188 CTCATATAATTTGCCAATTTCTGGAATTTCTATATAAATGAAGCAACCAACATGTGAGAC 247
 DB 6317 CTATATAGATTTGCCATTTCTGGACATTTCTATATAAATGAATCATATATATGTGGCT 6376
 QY 248 TTTGTGACTGGCTGCTTTTCCATTTAGCTTAGCATTTTAAAGGCTCA-TTATGTTACAGTAC 306
 DB 6377 TTTCTGACTAGCTTCTTTTCCATTTAGGGTATCATTTCAAGTTTCATTTATGTTAGCAT 6436
 QY 307 TTAGCAGTATTTCAITCTTTTTTATTTCTCAATGTTATTTCCATCTGTGGGTATCCCAT 366
 DB 6437 GTATCAGTGTTTTGTGTTTGTGTTTGTGAGCGGAGTTTGTCTCT----- 6482
 QY 367 TCATATTATAGACAGAGTTCTCACTGTCAACCAGCTGGAGTGCAGTGCACAATC 426
 DB 6483 -----TCITGCCAGGCTGGAGTGCAGTGGCGCAGTC 6514
 QY 427 ATAGCTCACTGTAACTCAAACTCTGGGCTCAAGTGATCTACTACTCAGCTCAGCTCCAGA 486
 DB 6515 TTGGCTCACTGTCAAACTCTGCTCTCTGGGTTCAAGTGATCTCTCCTCAGCTCAGCTCCGA 6574
 QY 487 GTAGCTAGGACTACAGGCACACACAGCCATACCTGGCTAAATTTTTTTTTTTTAAATTTCA 546
 DB 6575 ATAGCTGGGATTACAGGTGCACGCCACCATGCCAGCTAAATTTGTACTTTCAGTACACA 6634
 QY 547 TTTATGTATTC-----ATTTCTCTTTTCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 591
 DB 6635 TGGGGTTTCAACCATGTTGGCCAGGCTAGTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGA 6694
 QY 592 TAGGGTCTCACTTTTGTACCCAGGCTGGAGGCGAGTGGCATGGTGACAGCTGA--GCAGC 649
 DB 6695 GGGAGTCTCGTCTGTGCGCCAGGCTGGAGTGCAGTGGCGGATCTCGGCTCACTGCAAG 6754
 QY 650 CTGAGCTTCCTGGCTCAAGTGTATCTCTGCTCAGCTCAGCTCCCAAGTAGCTGGGACTACA 709
 DB 6755 CTGGCTCTCGGGTTTCAGGCCATTTCTCTGCTCAGCTCAGCTCCCGAATAGCTGGGACTACA 6814
 QY 710 AACACGTGTCAACCTCGCTGGCTGATTT-----TTTTTTCTTGAACAGGGTATCACT 763
 DB 6815 GGGCTCTGCCACCGCCCGGCTTAATTTTGTAGTATTTTGTAGAGACGGGTTTCAAC 8874
 QY 764 CTGTTCCCGAGGCTGGAGTACAGTGGCGTAAATAAGCTCACTGACAGCTCCCTCTCTGG 823
 DB 6875 GTGTTAGCCAGGATGTTCTCGATCTCTCACTCGTGATCCACCCGCTCGGCTCCCA 6934
 QY 824 GCTCAAGCAATCGCTGGCTCAGCTCAGCTAGTGGGACTACAGGCTTGTGGCACC 883
 DB 6935 AGTGTAGTATTACAGGCTTGAGCCACCGCGCCCGCCAGGCTAGTCTTGAATCC 6994
 QY 884 AGSCCCAGCTAACTTTTAAATAATGATTTTGTGTATAGAGAGGTTCTTGTATGTTGCTC 943
 DB 6995 TGACCTCTGTGATCTGCCCACTCTCAGCTCCCAAAAGTGTGGGATTAAGGATGAGCC 7054

944	QY	AGCGTGTATTTTATTTGTTGAGACAAGGTCCTCACTAGTTGCGCAGATGATCCCCCACTCC	1003
7055	Db	ACCATGCCCGCGTCGAGTGTTTTATTCCTTTTGTGTGCTGAATAGTATCTCATTTTATGG	7114
1004	QY	ACTTCCCAAGTGTCTCATCTTATCTGTTCATTAAGTCAGTTGAGACATTTAGTGTGTTT	1063
7115	Db	ACATAC-----TCATATTTTATGCACTTCATTTGTGCAGCTAAAGACATTTGGTAACTT	7168
1064	QY	CCACTTTTTCACCATATTGAATAATACTCCAGTGAATATTCATGT---ATACATATTGTGT	1120
7169	Db	CTACTTTTTCGATATTAAGTAATAATGCTTCGGAACATTCGTATACAAGATTTTAATGT	7228
1121	QY	GGGCAATATGTTTCATTTCTGTGGGTTTATATCTAGAGTGGAAATGCTGGATCCCGGG	1180
7229	Db	GGATCTATGTTTTCATTTTCTTGGGTATGTATCTCCTAGGACAGAAATGATGGGTCAATGT	7288
1181	QY	TAATATTTTGACAGGCAGAGTTCAGGGAGAAAAAATCTGGAAAAATGAAGCATGTTTAG	1240
7289	Db	GTGACTCTGTAGAAGCTTTTCAAGTCGCGTACTGTCTCCAGAGTGGCTGCATATTTTCC	7348
1241	QY	AAATCAGCAAGAGTCGAGGGGTTTTTTCGAGTGTTTATTTTATATTCGTGTGACAAATGTG	1300
7349	Db	ATTGATGTTGAA CATCTTTTCATGGGCTTATCGACGATTTGTATATCTTTGGAGAAATGC	7408
1301	QY	CAGTTTGATGAAGATACAAGTTTATCTPAAGTGAGAGTGAGAAATTAAGCTGGAAATAGG	1360
7409	Db	CCATTCGGTTCTTTTGCTGATTTTAAATTAATTTATCTTTTATTAATCTGAGTTGTGAAGCG	7468
1361	QY	CGTTTCAGAGTAAATCATGAAGACATTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAAC	1420
7469	Db	AAATCTTTTAGATAGTCAGATACAAGTCCCTTATCAGA-----TATAT	7511
1421	QY	AAAAATAAAAAATCACAATTTTTTTTTTTTTTTTTTTTGGAAAAAGAGCTTGTCTCTTCAAC	1480
7512	Db	TATTTGCAATATTTCTCCCATTTTGTGGGTGTCTTTCGACCTTCTTGTCTTTTCAA	7571
1481	QY	CTGGCTGGAGGCGAGTGGTGTGATCTCAGCTCACTGCAACTTTTGCCTCCCGGGTTCAAG	1540
7572	Db	TTG-----GAGTAAAAATATACATAACATAAAATTTAACTTTTGGGGTTGGGGCAAT	7625
1541	QY	CAATTCCTCTGCTTCAGCCTCCCAAGTAGCTGGGACTACAGGCATCTCCACCATGCCCA	1600
7626	Db	GGTTCACACTGTAAATCCAGCAATTTAGAGGGCAGGGTAGGAGGATGCTTTCAGCCCA	7685
1601	QY	GCTGATTTTTGTATTTTGTAGAGATGGCATTTCACTTTGTTGGCCAAAGCTGTCTCAA	1660
7686	Db	GGAGTTGAGACACAGCTCTGGGCAACATAGCAAGATCTGTCTCTACAAA-----	7734
1661	QY	ACTTTTTGCTGTATATTTGTGTAACTATGTTTCCTTTTGTCTGAGTAGGGCCCCCAGA	1720
7735	Db	-----ATATTTAAAAATTAGCTGGGCTGGCTGCACATCTCTAGTCCCAAG	7781
1721	QY	CCAAAAAATAAATCTTAGAATCCAAATCAGTGTGTTGGTTTGGACACATGTCATCTGAG	1780
7782	Db	TGAGGTGAGAAAAATCGCTTGAGTCCAGGAGGTTCAGGGCTCTAAAGTGAGCCGTGATCACAC	7841
1781	QY	AACCAAGCTGTACCGGGCTCAGGAGTAGAGGTGATCTCTGCTGAAAGAGAAATAGA	1840
7842	Db	CATGTCAGTCCAGCGTG-----ATGCGAGGTGAGACTCTGTATTTTTTAAAAAATAA	7894
1841	QY	ATGAAAAATATTCTCCGGCCAGGCGTGTGGCTCATGCCCTGTAATCCAGCACTTTGGGA	1900
7895	Db	AACAAAAACAAABAGAGCCAGGCGGGTGGCTCACGCCCTGTAAATCCAAGCACTTTGGG	7954
1901	QY	GGCCAAAGGCATGTGGATCACTTGAGTTCAGAGTTTCAAAACCGAGCTTGGCCACATGGTG	1960
7955	Db	GACCAAGCGCGGAGATCACTTGAGTCTGGAGGTTTCGAGACCAAGCCCTGACTACACGAG	8014
1961	QY	AAACCCGCTCTCTACTAAAAAATCAAAAAATTAGCTAAGTGTGGTGGCGCATGCTCTAA	2020
8015	Db	AAACCGTGCCTCAACTGAAATTC-----AAAAATTTTCCAGGTGTGGTGGCGCATGCTGTAA	8072
2021	QY	TCCAGCTACTCTGGAGGGGTGAGCGAGGAGAAATTTCTTGAACCCGGGAGGCGAGGTTGC	2080

Db	8073	TCCAGGCTATTCAGGAGTTTGGAGCAGGAATCGTTTGAACTGGAGCAGAGGTTC	8133
Qy	2081	AGTGAAGCGAGATCACACCACTGCACCTCCAGCCTCGGGGA-GAGAGCGAGACTTCCTCTC	2139
Db	8133	GGTGGCCAGATCGCACCATTCGACTCTAGCCTGGCAACAGAGCGAACTCCGCTCTC	8192
Qy	2140	AAAAAACAAAAACAAAAAGAA	2161
Db	8193	AAAAAAAAAAAAAGAAAAA	8214
RESULT 5			
ID	ABQ88096/C		
XX	ABQ88096 standard; cDNA; 92638 BP.		
XX	AC		
XX	ABQ88096;		
DT	18-SEP-2002 (first entry)		
XX	Human osteoblast differentiation related cDNA SEQ ID NO 3.		
DE			
XX	Human; osteoblast; stem cell differentiation; bone tissue deposition;		
KW	osteoporosis; osteopathic; ss.		
KX			
XX	Homo sapiens.		
OS			
XX	WO200250301-A2.		
FN			
XX	27-JUN-2002.		
PD			
XX	18-DEC-2001; 2001WO-US048276.		
PF			
XX	18-DEC-2000; 2000US-0255882P.		
PR			
PR	24-APR-2001; 2001US-0285691P.		
XX	(GENE-) GENE LOGIC INC.		
PA	(PROC) PROCTER & GAMBLE CO.		
PA			
XX	Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;		
PI	Mertz L;		
PI			
DR	WPI; 2002-557663/59.		
XX			
XX	Use of genes and their expression profiles associated with osteoblast		
PT	differentiation for screening modulators bone formation, for diagnosing		
PT	or treating e.g. osteoporosis, or as markers for the differentiation		
PT	process.		
XX			
PS	Claim 1; SEQ ID NO 3; 78pp + Sequence Listing; English.		
XX			
CC	The invention relates to genes and their expression profiles are used		
CC	for: (a) screening modulators of precursor stem cell differentiation into		
CC	osteoblasts, or bone tissue deposition; (b) diagnosing abnormal		
CC	deposition of bone tissue, abnormal rate of osteoblast formation or		
CC	osteoporosis; or (c) treating or monitoring treatment of the conditions		
CC	cited in (b), or monitoring the progression of bone tissue deposition.		
CC	Specific conditions include postmenopausal osteoporosis, glucocorticoid		
CC	osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-		
CC	induced abnormalities in bone formation or bone loss, conditions that		
CC	involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),		
CC	skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome		
CC	or fibrous dysplasia. The present sequence is that of an osteoblast		
CC	differentiation associated cDNA marker of the invention. Note: The		
CC	sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences		
XX			
XX	Sequence 92638 BP; 21460 A; 24572 C; 25235 G; 21371 T; 0 U; 0 Other;		
Qy	Query Match	9.58; Score 375.4; DB 6; Length 92638;	
XX	Best Local Similarity	55.44; Pred. No. 5.5e-67;	
XX	Matches 1046; Conservative	0; Mismatches 726; Indels 115; Gaps 12;	

QY 366 ATCAATATTTAGACAGAGTTCTCACTCTGTGCACCAGGCTGGAGTGGGACAAAT 425
Db 57931 AGCAATTTTTTTAGACAGAGTTCTCTCTGTGCACTCAGGCTGGAGTGGGCGAGT 57872
QY 426 CATAGCTCACTCACTCAAACTCCTGGGCTCAAGTGTATCTACTCAGGCTCCAG 485
Db 57871 CATAGCTCACTCAGGCTCGACCTTCAGGGCTCAAGCAATCTCTCTGCTCAACCTCCA 57812
QY 486 AGTAGCTAGGACTACAGGCACACAGGCCATACCTGGCTAAATTTTTTTTTTAATTTCA 545
Db 57811 AGTAGCTGGACTACAGGCA-GCAACACTATGGCTGGCTAAATTTTACATTT- 57761
QY 546 TTTTANGATTCATTTTCTTTCTTTTGTGTTGTTGTTGTTGTTGATAGGCTCAGTTT 605
Db 57760 -----TTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 57714
QY 606 GTTACCCAGGCTGGAGGAGTGGCATGGTGACAGCTGA--GCAGCCTTTGACTTCTCTGG 663
Db 57713 GTCACCAGGCTGAAGTACAGTGTGGGATCTTGGCTCACTGCCAGCTCCACCTCCGGG 57654
QY 664 CTCAGTGTATCTCTGCTCAGCTTCCCAAGTAGTGGGACTCAAAACAGTGTCAACA 723
Db 57653 TTCAAGCCATCTCTCTGCTCAGCTTCCGAGTGTGGGACTACAGGCGGCCACCATCA 57594
QY 724 TGCTGGCTGA-----TATTTTTTTTCTTTGAAACAGGCTATCACTCTGTGGCCAGGCTG 778
Db 57593 CGCCGAGCTAACTTTTGTATTTTGTAGTAGACGGGTTTCAACATGTTAGCCAGGATG 57534
QY 779 GAGTACAGTGGGCTAATAATAGCTCACTGCAGCCTCCCTCC-----TGGGCTCAAGCAATC 835
Db 57533 GTCTCGATCTCTGACCTCGCGATCGCCCGCTCGGCTCCCAAGTGTCTGGATTACA 57474
QY 836 CGCTGGCTCAGCATCTGAGTAGTGGGACTACAGGCTGTGGCCACGAGGCCAGCTAA 895
Db 57473 GGAGTGAGCCACTATGCGCGGCTAAGTTTAAATTTTTCATAGAGCAAGGCTCTCACTA 57414
QY 896 GTTTTAAATAAATGATTTTTGTTATAGAGGAGTCTTGTCTATGTTGCTCAGGCTGATTTT 955
Db 57413 TTGTTGCCAGGCTGTCTCAAACTTACGTGATCTCTCTACCTTAGCTTCCCAAGGGCT 57354
QY 956 TATTTGTGAGACAAGTCTCACTATGTTGCCATGATCCCGCCACCTCCACTTCCCAAGT 1015
Db 57353 GGGATTACAGGTTGGTCTCCCTATTATTAACACTTAGTATGTCATATTTGTCAAT 57294
QY 1016 GCTCATCTTATCTGTTATAGTGTGAGTGTGACAGACATTTAGTGTGTTTCCACTTTTGC 1075
Db 57293 GAATCATTTAATAAACTCCATATCTTTTATTTACGATTTTACTCAGTTTCCCGCTAATGTC 57234
QY 1076 CATTATGAATAATCTCCAGTGAATATTCATGTATACATTTGTGGGCATATGTTTCA 1135
Db 57233 CTTTTCCTCAGATCCCATTTCAATCATCTGACTCTTGGATGTGATGTTTCTCA 57174
QY 1136 TTTCTGTTGGGTT----- 1149
Db 57173 GACCTTCCCTGTTTTTGTATGACCTTACAGTTTTTGGGAGTACCGGTAGGCATATGGTAG 57114
QY 1150 -----ATATCTAGAGTGGAAATGCTGGATCCCGGGTAATATTTTGACAGGAGATT 1202
Db 57113 AATCCCTCTATCAGGATTTGCTGATGTTTCTCATGATTTAGACTGAGGTTCTGGGTT 57054
QY 1203 CAGGGGAAGAAACTTGGGAAATGAAGCATGTTTAGAAATC--AGCAAGAGTGCAGG 1260
Db 57053 TTGGAGAGAGACCAAGAGTCAAGTGGCCATTTTGGCCACATCTTATGAAGGTTCTATGC 56994
QY 1261 GTTTTTTCGGAGTTTATTTATTTATTTCTGTGCAAAATGTCAGTTTGTGAAGATACAG 1320
Db 56993 TGACTTTGATCACTGACTGAGTAAACTTTGCCAGACTTCTCCATGTGAAAGTACTCT 56934
QY 1321 TTATATAAGTAGAAGTGAAGTAAAGGCTGAATAGGGCGTTTCAGATTAATCATGA 1380
Db 56933 TTTTCTTCTTCTGTACTGGACCTTTTGAAGGAAGTCACTGTGTGCAGCCCATTTTA 56874

QY 1381 AGCACTTTGAATACCAAAATTAAGAGCTTGGCTCTTAACAACAAATAATAAAAAATCAAA 1440
Db 56873 AGAATTTGGGA-ATTATGTTCTACTCTCTTAAGGCGAGAAATATCTACAGGTTTTTGT 56815
QY 1441 TTTTTTTTTTTTTTTTGAAGAAAGAGTCTTGTCTTTTCAACCTGGCTGGAGGCGAGTGGTG 1500
Db 56814 TTTTTTTTTTTCTTTTGAATGGAGTCTCGTCTGTGTCGCCAGGCTGGAGTGCAGTGGCA 56755
QY 1501 TGAATCTAGCTCAGTCAAACTTTGCGCTCCCGGGTTCAAGCAATTTCTCTGTTTCAAGCT 1560
Db 56754 CAATCTGCTCAGTCAAAAGCTCTGCTCTCCCGGGTTCAACAATTTCTCTGCTCAGCT 56695
QY 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCCAACCATGCCAGCTAATTTTGTATTTTAG 1620
Db 56694 CCCGAGTAGCTGGGACTACAGGCGGCCGCCACCATGTCCAGCTAATTTTGTATTTTAG 56635
QY 1621 TAGAGATGGGATTTCACTTTGTTGGCCAAAGTGTGTTCTCAAACTTTTTCG----- 1669
Db 56634 TAGAGACGGGTTTTCACTTTGTTAGCCAGGATGGTCTCGACTCTCGACTTGTGATCCG 56575
QY 1670 ----TGTCAATAATTTGTTGTAATACTATTGTTCTTTTGTCTGAGTAGGGCCCCCAGACCAA 1725
Db 56574 CCGGCTCGAATCTCAAAAGTGTGGGATTAACAGCGTGTAGCCACCCGCGCAGTTT 56515
QY 1726 AAAAAATAAATCTTAGAATCCAAATCAGTGTGTGTTGTTGACCACTGTCACTTTGAGAACCA 1785
Db 56514 ATTTGGAATTTCTCTGCAATGGGAGATTTGCTTATTTCTCCCTTATTTATTCATTCAT 56455
QY 1786 CAGTGTACCGAGGCTCAGGAGTAGAGTGTATCTCTGCTCGAAAGAGAAATAGAAATGAA 1845
Db 56454 TTATATCAGCATGGACTTACAGGTATTTTATTTTACTTTTGGGATATGATCTCAAACTAC 56395
QY 1846 AATATTTCTC-----GGGCCAGGCTGTGTGCTCATGCTGTAATCCAGCACTTT 1896
Db 56394 TTAATTTTGTGCCAAGTGTGCCAGGCTGTGTGCTCACCTGTAGTCTCCAGCACTTT 56335
QY 1897 GGGAGGCCAAGGCTATGTGGATCACCTGAGGTCAGAGTTCAAAACCAAGCTTGGCCAAAT 1956
Db 56334 GGGAGGCCAGGCGAGGCGGATCAC--GAGGTCAGGAGATCAAGACCATCTTGGCTTAACAC 56277
QY 1957 GGTGAAACCCCGTCTCTACTAAAAATACAAAATTAAGTGTGTGGGCGATGCT 2016
Db 56276 AGTGAACCTTGTCTCTATTAATAAATACAAAAAATAGCCAGCATGTTGGCGGCGCT 56217
QY 2017 GTATCCAGCTACTTGGGCGGTGAGGAGAGAAATTTCTTGAAACCGGAGGCGAGG 2076
Db 56216 GTAGTCCAGCTACTTGGGCGGTGAGGAGAGAAATGTTGTGAACCCCGGAGGCGGAGC 56157
QY 2077 TTGCAAGTGAAGCGAGATCACACCACTGCACTCCAGCCTGGGGAGAGAGGAGCTTCT 2136
Db 56156 TTGCAAGTGAAGCGAGATCGGCCACTGC-CTCCAGCCTGGGTGACAGAGCGAGCTCCAT 56098
QY 2137 CTCAAAAAACAATAAACAAGAAATTT 2163
Db 56097 CTCAAAAAGTAAAAATAAATAAATTT 56071

RESULT 6

AAK69044/c

ID AAK69044 standard; DNA; 5491 BP.

XX AAK69044;

AC AAK69044;

XX 06-NOV-2001 (first entry)

XX

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23856.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX

FN WO200157182-A2.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;

DR WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX Disclosure; SEQ ID NO 23856; 3071pp + Sequence Listing; English.
PS
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients' own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK67694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 5491 BP; 1662 A; 1226 C; 1439 G; 1164 T; 0 U; 0 Other;
Query Match 9.3%; Score 368.8; DB 4; Length 5491;
Best Local Similarity 56.0%; Pred. No. 6.5e-66;
Matches 1023; Conservative 0; Mismatches 627; Indels 178; Gaps 10;
QY 361 CCCATCATATATTAGAGACAGGTTCTCATCTGTCCACCGGTGGAGTGCAGTGGC 420
DB 3658 CCTAATTAATTTTGTGAGCGGTCTCTCACTGTGCGCCCAAGTGGAGTGAATGAC 3599
QY 421 ACAATCATAGTCACGTAACTCAATCTCTGGGCTCAAGTGATCTACTACCTCAGC 480
DB 3598 TGAATCTTGGCTCCTACGTCAACCTCCACCTACCGGGTTCAAGTGATCTCTCCGCTCAGCC 3539
QY 481 TCCAGAGTAGCTAGGACTACAGGCACACACAGCCATACCTGGCTAAATTTTTTTTAAAT 540
DB 3538 TCCCAAGTAGCTGGGACTATAGCGGGACCATCATCGCTTGGCTAATTTTGTATT--- 3483
QY 541 TTTCATTTTAAATGATTCATTTCTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
DB 3482 -----TTTAGTAAAGACAAAGGTTT 3463
QY 601 ACTTTGTTTACCCAGGCTGGAGGCGAGTGGCATGGTGACAGCTGAGCGCTTGTACTTCT 660
DB 3462 GCCTGTTGGCCAGGCTGGT-----CTTGAATCAT 3432
QY 661 GGGCTCAAGTATCTCTCTGCTCAGCCCTCCCAAGTAGCTGGGACTACAAACACGTTCA 720
DB 3431 GACCTCAGGTGATCCGCTCATCATCAGCTCCCAAGTGTGGGATTAACAGGCGTGAGCCA 3372
QY 721 CCATGCTGGCTGATA-----TTTTTTTCTTGAACAGGGTATCA 761
DB 3371 CTGTGCTGTGCCATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 3312
QY 762 CTCTGTTGCCAGGCTGGAGTACAGTGGCGTAAATATAGCTCACTGACGCTCC-CTTCC 820
DB 3311 CTCTGTTGCCAGGCTGGAGTACAGTGGCGTAAATATAGCTCACTGAAACCTCCATCTCC 3252
QY 821 TGGGCTCAAGCAATCCGCTGGCTCAGCATCTCTGAGTGTGGGACTACAGCTTGTGTC 880
DB 3251 TAGGTTCAAGCTATTCTCTGCTTCAAGCTCCCAAGTAGTGGGATTAACAGGCGCCACC 3192
QY 881 ACCAGGCCAGCTAAGTTTAAATATGATTTTGTATAGAGGAGGCTCTGCTATGTTG 940
DB 3191 ACCACACCCGGCTAATTTTGT-----ATTTTATGATGAGGATTTCAACCATGTTG 3138
QY 941 CTCAGGCTGTATTTTATTTGTTGAGACAAAGGCTCTCATGTGTCATGTGTCATGATCCCCCACC 1000

DB 3137 GCCAGGCGGCTCACCAACTCTCGGCTCAA-----ATGATCCACCTGCC 3094
QY 1001 TCCACTTCCCAAAGTGTCTCATCTTATCTTGTTCATTTAGTCAGTTGACAGACATTTAGGTTG 1060
DB 3093 TCAGCTTCCCAAAGTGTCTGGGATTAACAGCGCTGATCCACCGCGCCCGCTGAATGTTTC 3034
QY 1061 TTTCCACTTTTGTGACCAATATGAATTAATCTCCAGTGAATATTCATGTATATACATTTGTGT 1120
DB 3033 TTTCTTTCTTTTATTTTCTTTTCTTTTATAGACATAGTCTTGTCTGTTTACCAATGCT 2974
QY 1121 GGGCATATGTTTTTCTATTTCTGTTGTTTATATCTAGGAGTGAATTCCTGATCCCGGG 1180
DB 2973 GGAGTACGGTGGCGAGATCAGACTCAGCTACACTTGAATTCCTGGATTCAGGAATTC 2914
QY 1181 TAAATTTTGAAGCAGAGATTCAGGGGAAGAAAACCTTGGGAAATGAAGCATGTTTGTAG 1240
DB 2913 TCCCATCTCAGCTCTCTGAGT-----AGCTAAG 2886
QY 1241 AAATCAGCAAGATGACAGGGTTTTTCGGAGTTTTTATTTATTTATTTATTTATTTATTTAT 1300
DB 2885 GACTACAGGTCGTACCACTCTCTGCTGTTTCTGAATATCTGTTTATATAAATTC 2826
QY 1301 CAGTTTGTATGAAGATACAAAGTTTACTAAGTGAGAGTGAGAAATTAAGCTGGAATAGG 1360
DB 2825 CTATTTGTTTCAATACCTCTCTCAAGATATCTATTTATAGTCTGTCTCCGATTT 2766
QY 1361 CGTTCAGAGTAAATCATGAGACATTTTGAATACCAAAATTAAGAGGCTTGGCTGTAAAC 1420
DB 2765 TCTTGTGATCTTGCAACTGAAATGTTTCTGTATCTCTGGGACTGAATTTCTT 2712
QY 1421 AAATAATAAAATACAAATTTTTTTTTTTTTTTTGTGAGAGAGTCTTGTCTCTTACC 1480
DB 2711 -----TTTCTTTT-----TTTTTGTAGACAGATPACTTGTCTGTCACC 2665
QY 1481 CTGGCTGAGGCGAGTGTGTGATCTCAGCTCACTGCAACTTTTGGCTCCCGGGTCTCAAG 1540
DB 2664 CAGGCTAGATTCAGTGTGCTTGAATCTAGCTCAITGCAACCTCCGCTCTCTGGGTTCAAG 2605
QY 1541 CAATTTCTCTGCTCAGCTCCCAAGTAGCTGGGACTACAGGACTTCCACAGCATGCCCA 1600
DB 2604 CAATTTCTGCTCAGCTCCCAAGTAGCTGGAATTAACAGATGCATGCCACACGCCCC 2545
QY 1601 GCTGATTTTCTGATTTTGTAGTAGAGTGGGATTTACATTTGTTGGCCAGCTGTGCTCAA 1660
DB 2544 ACTAATTTTGTATTTTGTAGTAGGACTGGGTTTCCCATGTTGGCCAGGCTGTTGT 2485
QY 1661 ACTTTTGTCTGTCAATAATTTGT-----TGTAACATTTTGTCTTGTGTGAGGTAG 1710
DB 2484 ACTCTGACCTCAAGTGTATCTGCCGCTCTCGGCTCCCAAGTGTGGGATTTATAGGAT 2425
QY 1711 GGCCCCCAGACCAAAAAAATAATCTTAGAATCAAAATCAGTGTGTTGTTGACCACT 1770
DB 2424 GAGCCACCATGCCAGCCATGAGACTGAATTTTCTGTTGTTTGTGTTTGTGGGCTCTATTT 2365
QY 1771 GTCACTTCAGAACCAACAGTGTGACAGGCTCTCAGGAGTAGAGTGTATCTCTGCTCGAAA 1830
DB 2364 TTATGTTAGGACTGTTTCTCCAGCGTTGGTGTCTTTGGCTCTATATTCTCTACA 2305
QY 1831 GAGAAATGAATGAATAATTTCTCGGCGCCAGGCTGTGGCTCATGCTGTAAATCCGAG 1890
DB 2304 AGCAAGGCACTAAAAAGCTGGTTGAAGGCCAGACACAGCGGCTCACGCTGTAAATCCGAG 2245
QY 1891 CACTTTGGAGGCCAAGGCATGTGATCACTCAGTGTGAGGTTCAAAACACGCTCGC 1950
DB 2244 CACTTTGGAGGCTGAGGCGAGGATCA--TGAGGTGAGGAGTGGAGGCTCTGATCTCTGAC 2187
QY 1951 CAACATGTGAAACCCCGCTCTCTACTAAAAATACAAAAAATTAGCTAAGTGTGGGGCGC 2010
DB 2186 TAACTCGGTGCAACCTGTCTCTACTAAAAATAAATAAATAATAGCAGTGTGGTGGCGG 2127
QY 2011 ATGCTCTGAATCCCAAGCTACTTGGAGGAGTGGAGGAGGATTTCTTGAACCCGGGAGG 2070
DB 2126 GTGCTGTAGTCCCAAGCTACTTGGAGGCTGAGGCGAGGAGTGTATGAGCCAGGAGG 2067

PR	01-SEP-2000;	2000US-02239440P
PR	01-SEP-2000;	2000US-02239450P
PR	05-SEP-2000;	2000US-02239455P
PR	05-SEP-2000;	2000US-02239465P
PR	06-SEP-2000;	2000US-02295133P
PR	06-SEP-2000;	2000US-0230437P
PR	06-SEP-2000;	2000US-0230438P
PR	08-SEP-2000;	2000US-0231242P
PR	08-SEP-2000;	2000US-0231243P
PR	08-SEP-2000;	2000US-0231244P
PR	08-SEP-2000;	2000US-0231413P
PR	08-SEP-2000;	2000US-0231414P
PR	08-SEP-2000;	2000US-0232080P
PR	08-SEP-2000;	2000US-0232081P
PR	12-SEP-2000;	2000US-0231968P
PR	14-SEP-2000;	2000US-0233297P
PR	14-SEP-2000;	2000US-0233298P
PR	14-SEP-2000;	2000US-0233299P
PR	14-SEP-2000;	2000US-0233400P
PR	14-SEP-2000;	2000US-0233401P
PR	14-SEP-2000;	2000US-0233603P
PR	14-SEP-2000;	2000US-0233604P
PR	14-SEP-2000;	2000US-0233605P
PR	21-SEP-2000;	2000US-0234223P
PR	21-SEP-2000;	2000US-0234274P
PR	23-SEP-2000;	2000US-0236327P
PR	23-SEP-2000;	2000US-0236328P
PR	23-SEP-2000;	2000US-0236367P
PR	23-SEP-2000;	2000US-0236368P
PR	23-SEP-2000;	2000US-0236369P
PR	23-SEP-2000;	2000US-0236370P
PR	02-OCT-2000;	2000US-0236802P
PR	02-OCT-2000;	2000US-0236802P
PR	02-OCT-2000;	2000US-0237037P
PR	02-OCT-2000;	2000US-0237038P
PR	02-OCT-2000;	2000US-0237039P
PR	02-OCT-2000;	2000US-0237040P
PR	13-OCT-2000;	2000US-0239935P
PR	13-OCT-2000;	2000US-0239937P
PR	20-OCT-2000;	2000US-0240936P
PR	20-OCT-2000;	2000US-0241221P
PR	20-OCT-2000;	2000US-0241785P
PR	20-OCT-2000;	2000US-0241786P
PR	20-OCT-2000;	2000US-0241787P
PR	20-OCT-2000;	2000US-0241808P
PR	20-OCT-2000;	2000US-0241809P
PR	20-OCT-2000;	2000US-0241826P
PR	01-NOV-2000;	2000US-0244617P
PR	08-NOV-2000;	2000US-0246474P
PR	08-NOV-2000;	2000US-0246475P
PR	08-NOV-2000;	2000US-0246476P
PR	08-NOV-2000;	2000US-0246477P
PR	08-NOV-2000;	2000US-0246478P
PR	08-NOV-2000;	2000US-0246523P
PR	08-NOV-2000;	2000US-0246524P
PR	08-NOV-2000;	2000US-0246525P
PR	08-NOV-2000;	2000US-0246526P
PR	08-NOV-2000;	2000US-0246527P
PR	08-NOV-2000;	2000US-0246528P
PR	08-NOV-2000;	2000US-0246532P
PR	08-NOV-2000;	2000US-0246609P
PR	08-NOV-2000;	2000US-0246610P
PR	08-NOV-2000;	2000US-0246611P
PR	08-NOV-2000;	2000US-0246613P
PR	17-NOV-2000;	2000US-0249207P
PR	17-NOV-2000;	2000US-0249208P
PR	17-NOV-2000;	2000US-0249209P
PR	17-NOV-2000;	2000US-0249210P
PR	17-NOV-2000;	2000US-0249211P
PR	17-NOV-2000;	2000US-0249212P
PR	17-NOV-2000;	2000US-0249213P
PR	17-NOV-2000;	2000US-0249214P

```

RESULT 7
AAS42015/C
ID   AAS42015 standard; DNA; 9742 BP.
XX
XX
XX   AAS42015;
XX
XX
XX   17-DEC-2001 (first entry)
XX
XX
XX   Genomic sequence #331 encoding novel human enzyme polypeptide.
DE

```

Genomic sequence #331 encoding novel human enzyme polypeptide.

Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase; ligase; hyperproliferative disorder; immunodeficiency disorder; autoimmune disorder; neurological disorder; metabolic disorder; inflammatory disorder; cardiovascular disorder; reproductive disorder; blood-related disorder; infectious disorder; degenerative disorder; anti arthritic; nephrotropic; anticonvulsant; ds.

OS Homo sapiens.

PN WO200155301-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US001239.

PR 31-JAN-2000; 2000US-0179065P.

PR 24-FEB-2000; 2000US-0184664P.

PR 16-MAR-2000; 2000US-0189874P.

PR 18-APR-2000; 2000US-0198123P.

PR 07-JUN-2000; 2000US-0209467P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217496P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225758P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226868P.

PR 23-AUG-2000; 2000US-0227009P.

PR 01-SEP-2000; 2000US-0229287P.

PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465566/50.
XX Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.
XX Disclosure; SEQ ID NO 2141; 1180pp; English.
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders, including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAS41685-AA42192 represent
CC DNA sequences encoding for the novel human enzyme polypeptides of the
CC invention. Note: the sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 9742 BP; 2382 A; 2564 C; 2516 G; 2280 T; 0 U; 0 Other;
Query Match 9.3%; Score 366.4; DB 4; Length 9742;
Best Local Similarity 56.3%; Pred. No. 2.3e-65;
Matches 1043; Conservative 0; Mismatches 631; Indels 178; Gaps 12;
370 TATTATTAGAGACAGGTTCTCACTC-TGTCAACCCAGGCTGGAGTGGCAACAATCAT 428
6351 TTTTCTTTGAGATGAGATTTCATCTTGTGGCCAGGCTGGAGTGTGTGGCCATCTTT 6292
429 AGCTCACTGTAACTCAAACTCCCTGGGCTCAAGTGATCTTACTACTCAGCCTCCAGAGT 488
6291 GGCTCAATTCGAACTCCGGCTCCGGGGTTCAAGCAATTCCTCCCTCAGCCTCCCTGAGT 6232
489 AGCTAGAGCTACAGGACACACAGCCGATCTGGCTAAATTTTTTTTAAATTTTCAATTT 548

Db 6231 AGCTGGGTTTACAGCATGTGCCACACGCCCGGCTAAATTTTGTATTTTAA----- 6181
QY 549 TATGTATTTCATTTTCTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 608
Db 6180 -----GTAGAGGCGGGTCTTCCATGTT 6157
QY 609 ACCCAGGCTGGAGGCGCAGTGGCATGGTGACAGCTGAGCAGGCTTGACTTCCTGGGCTCAA 668
Db 6156 GGTCAAGGCTGGT-----CTTGAACCTCCGCACTCAG 6126
QY 669 GTGATCTCTCGCTCAGCCTCCCAAGTAGCTGGGACTACAAACACGCTGTACCATGCT 728
Db 6125 GTGATCCCGCCCGCTCGGCTTCCCAAGTGTGGGATTACAGGCTGAGCCACACGCTC 6066
QY 729 GGCTGATATTTTTTTTCTTTGAAAACAGGGGTATCACTCTGTGCCCCAGGCTGGAGTACAGTG 788
Db 6065 GGCCTTTTATTTTGTGAGATGGAGTTTGGCTCTGTCAACCCAGACTGGAGTCAAGTG 6006
QY 789 GGGTAAATATAGCTCACTGACGCTC-CCCTCTGGGCTCAAGCAATCCGCTGGGCTCAG 847
Db 6005 GTGCAGTCTCAACTTACTGCAACCTCTGCTCCTAGGTTCAAGCAATTTCTCATGCTCAG 5946
QY 848 CATCTGTAGTGGGCTGAGGCTGTGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 907
Db 5945 CTTCCCAAGTAGATGGGACTCAGGCATGTGCCACACGCCCTGGCTGCTGCTGCTGCTGCTGCT 5890
QY 908 GATTTTGGTATAGAGAGGCTCTTCTGCTATGTTGCTCAGGCTGTATTTTATTTTGTAGAC 967
Db 5889 TTTTGTAGTAGAGACGTTGGTTTCTGCTCTGTGTGGTAGGCTGGTGTGAAACGCTGACCT 5830
QY 968 AAGGCTCATTATGTTGCCATGATCCCCCAGCTCCTCCTCCCAAGTGTCTCATCTTATC 1027
Db 5829 CAAG-----TGATCCGCCACCTCCGCTTCCCAAGTGTGGGATTATA 5786
QY 1028 TGTTCATTAGTCAGTTGACAGACATTTAGGTTGTTTCCACTTTTGAACCATTAATAA 1087
Db 5785 GCGGTGAGCCAGGCACTGAGTGTGAATATTTTAAAGGCCAGTCAATGATCTAAT 5726
QY 1088 TACTCCAGTGAATATTCATGTATATCATTTGTGTGGGCATATGTTTTCATTTCTGTGGGT 1147
Db 5725 GTCTCTGGTAAATTCATGACCGCCCTTGGTTTATGACAGCGTTCTAGCCATAGCC 5666
QY 1148 TTATATCTAGAGTGGAAATGCTGATCCCGGTTATATTTGACAGGCGAGATTGAGG 1207
Db 5665 CATTTTACAGATAGGACCCAGTGGAGTCTTGGGAA----- 5629
QY 1208 GAAGAAAACTTTGGGAAATGAAGCATGTTTGAATCAGCAAG-AGTGCAGGGGTTTT 1266
Db 5628 -TAGAGATTTAGGAACTTGTACGTTGACAGCCAGGAAATATAGCAGCTAAGAA 5570
QY 1267 CGAGTTTATTTTATTTATTTCTGTTGACAAATGTGCAAGTTTGTATGAGATCAAGTTATAC 1326
Db 5569 AAGTCTTGCAGCTAGGCACTGGAGCTCACACCTGTAATCTCAGTGTCTTTGGGAGGCTGAG 5510
QY 1327 TAAGTGAGAAGTGAGAAATTAAGCTGGAATAGGGCTTCAGAGTAAATCATGAAGCACT 1386
Db 5509 GCAGAGAGATCGCTGAGGCCAGGAGTTTGAAGACCGCTGGGAAACATAGGAGACCT 5450
QY 1387 TTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATAATAAAATACAAATTTTT 1446
Db 5449 ATCTCTATGAAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5390
QY 1447 TTTTGTGAAAGAGTCTTGTCTTTTACCTTGGCTGGGCGGCTGGTGTGATCT 1506
Db 5389 TTTTGTGAGACAGAAATCTCACTCTGTGCCAGGCTGGAGTGAATGGTGTATCT 5330
QY 1507 CAGCTCAGTGCACTTTTCTGCTCCCGGTTCAAGCAATTTCTCTGCTCAGCCTCCCAAG 1566
Db 5329 CAGCTCACCACAACTCCCGCTCTCAGGTTCAAGTGAATTTCTCTCCTCAGCCTCCTGAG 5270
QY 1567 TAGCTGGGACTACAGGCACTTCCCACTGAGTGTGAGTGTGATTTTGTATTTTGTAGTAGA 1626
Db 5269 TAGTGGGACTACAGCGGGTGCACCATGTCCGGCTAATTTTGTATTTTGTAGTAGA 5210


```
Db      20447 ATCATACGGCGCGCGGTGGCTACGCTGTAATCCAGCACTTTGGAGGCGGAGG 20506
QY      1909 CATGTGGATCACTGAGGTGAGAGTTCAAAACGAGCCCTGGCCACATGTTGAACCCCG 1968
Db      20507 TGGGCGGATCAC--GAGGTGAGGATCGAGACCATCTCTGGCTAAACGTTGAACCCCG 20564
QY      1969 TCTCTACTAAATAACAAAAATTAGCTAAGTGTGTGGCGCATGCTGTAATCCAGCT 2028
Db      20565 TCTCTACTAAATAACAAAAATTAGCCGGGCTAGTGGCGGCTGCTGCTCCAGCT 20624
QY      2029 ACTTGGGAGGTGAGGAGGAGAAATTTCTTGAACCGGAGGAGGAGGTTGCAAGTAAAGC 2088
Db      20625 ACTCGGAGGCTGAGGCAGAGAGATGGCTGAACCGGAGGCGGAGCTTGCAGTGAGCC 20684
QY      2089 GAGATCACACCACTGCATCCAGCTCGGGGAGGAGGCGAGACTTCTCTCAAAAAACA 2148
Db      20685 AAGATCGCCCACTGCCCTCCAGCTGGCGCGACAGAGCGAGACTCGTCTCAAAAAAAA 20744
QY      2149 AAAAAACAAAGAAATTAAGCAAAATTAGACATTCAGAGAGAA 2189
Db      20745 AAAAAAATAAAAAAAGGAATCATCTTGAAGAGGAA 20785
```

RESULT 9

```
AAD14749
ID AAD14749 standard; DNA; 12394 BP.
```

```
AC AAD14749;
```

```
XX 01-NOV-2001 (first entry)
```

```
DE Human glycogen synthase kinase 3 alpha genomic DNA.
```

```
KW Human; glycogen synthase kinase 3 alpha; antidiabetic; cytostatic;
KW antisense therapy; diabetes; hyperproliferative disorder; inflammation;
KW neurological disorder; tumour; haematopoietic disorder; infection;
KW hyperproliferative disorder; developmental disorder; ds.
```

```
OS Homo sapiens.
```

```
XX Key Location/Qualifiers
```

```
FT CDS 115..11786
FT     /tag= a
FT     /product= "Human glycogen synthase kinase 3 alpha"
FT exon 115..397
FT     /tag= b
FT intron 398..2437
FT     /tag= c
FT exon 2438..2625
FT intron 2626..5638
FT exon 5639..5722
FT intron 5723..5863
FT exon 5864..5974
FT intron 5975..7901
FT exon 7902..8032
FT intron 8033..8120
FT exon 8121..8227
FT intron 8228..9196
FT exon 9197..9294
FT intron 9295..9374
FT     /tag= o
```

```
FT exon 9375..9470
FT     /tag= p
FT intron 9471..9897
FT     /tag= q
FT exon 9898..10084
FT     /tag= r
FT intron 10085..10430
FT     /tag= s
FT exon 10431..10523
FT     /tag= t
FT intron 10524..11712
FT     /tag= u
FT exon 11713..11786
FT     /tag= v
XX WO200152865-A1.
XX 26-JUL-2001.
XX 16-JAN-2001; 2001WO-US001411.
XX 21-JAN-2000; 2000US-00488856.
XX (ISIS-) ISIS PHARM INC.
XX Monia BP, McKay R, Butler MM, Wyatt JR;
XX WPI: 2001-442247/47.
XX P-PSDB; AAE06521.
XX Antisense compound 8 to 30 nucleobases in length comprising a compound
XX that is targeted to a nucleic acid molecule encoding glycogen synthase
XX kinase 3 alpha, useful for the treatment of e.g. diabetes and
XX hyperproliferative disorders.
XX Example 15; Page 92-100; 115pp; English.
XX The invention relates to an antisense compound 8 to 30 nucleobases in
XX length targeted to a nucleic acid encoding glycogen synthase kinase 3
XX alpha. The antisense compound specifically hybridises with and inhibits
XX the expression of glycogen synthase kinase 3 alpha. The antisense
XX compound is useful for the treatment of a diseases associated with
XX glycogen synthase kinase 3 alpha such as diabetes, a neurological
XX disorder, a haematopoietic disorder, a hyperproliferative disorder or a
XX developmental disorder. The antisense compounds may also be used
XX prophylactically to prevent or delay infection, inflammation or tumour
XX formation. The present genomic DNA sequence encodes human glycogen
XX synthase kinase 3 alpha protein
XX Sequence 12394 BP; 2780 A; 3043 C; 3547 G; 3024 T; 0 U; 0 Other;
SQ Query Match 8.8%; Score 349.8; DB 4; Length 12394;
Best Local Similarity 55.2%; Pred No. 6.4e-62;
Matches 1016; Conservative 0; Mismatches 742; Indels 81; Gaps 14;
QY 370 TATTATTAGACAGAGTTCTCACTCTGTCACCAGCTGGAGTGAGTGGCACATCATTA 429
Db 3604 TTTTITTTTGGATGGAGTCTTACTCTGTCACCAGGCTGGAGTGAGTGGCACATCTTG 3663
QY 430 GCTCACTGTAACCTCAAACTCTGGGCTCAAGTATCTTACTACCTAGCCTCCAGAGTA 489
Db 3664 GCCCGCTCAACCTCTGCCTCCAGGTTCAAGCGGATTTTGTGCTCAACACCTGAGTA 3723
QY 490 GCTAGGACTCAGGCGACACACAGCCATACCTGGCTAAATTTTTTTTTTAAATTTTCATTT 549
Db 3724 GCTGGATTGAGGCGATGGCCACTATGCTGGCTAAATTTTTTTTGTGTAATT--TTTA 3781
QY 550 ATGTATTCAATTTTCTTTCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTA 609
Db 3782 GTAGAGACAGAGGTTTGGCCATGTTGGCCAGGCTGTCTCGAACTCCTGACCTCAAGTAAT 3841
QY 610 CCCAGGCTGGAGGCGAGTGCGGATGGTGCACACTGAGCAGCCTTGACTTCTCTGGCTCAAG 669
```

Db 3842 CCGCCCCCACTCGACCTCCAAAGTCTGGTATTACAGGCATGA-----GCCAGCAT 3894
Qy 670 TGATTCCTCTCGCTCAGCTCCCAAGTAGCTGGACATACAAAACACAGTGTCA-CCATGCCT 728
Db 3895 ACCTAGCCCTGATTTTCAAGACAACTGAAACCTGGATTAGATGTGAATCTTTTTT 3954
Qy 729 GGCTGATATTTTTTTCTTGAACACAGGGTATCATCTGTGTGCCAGGCTGGAGTACAGTG 788
Db 3955 TTTTTTTTTTTTTTTTGGAGACGGAGTCTCATGTGTGTCACCCAGGCTGGAGTGTGGTG 4014
Qy 789 GCGTAATAATAGTCACTGACAGCTCC-CTCCTGGGCTCAAGCAATCCGCTGCTCAG 847
Db 4015 GCGTATCTCGGTCATGCAACTCCGCTGCGGGTCAAGGATCTTCTGCTCAG 4074
Qy 848 CATCCTGAGTAGCTGGGACTACAGGCTGTGSCACAGGCGCCAGCTAAGTTTTTAAAAAT 907
Db 4075 CTTCCCTAGTAGCTGGGACTACAGCGGTGTCACCACTCTCGGCTAATTTTTGTATTT 4134
Qy 908 GATTTTGGTATAGAGAGGCTTGTCTATGTGTCTCAGGCTGTATTTTATTTGTTG--- 963
Db 4135 CTAGTAGGACGGGGTTTACCGAGTTAGCCAGATGGTCTCTATTTTTTTTTTTTTTT 4194
Qy 964 ----AGACAGGCTCTACTATGTTGCCATGATCCCCCACTCCACTTCCCAAGTCTC 1019
Db 4195 TTTAAGACAGATCTGTTCTGTCTCACTAAGCTGGAGTGCAGTGGTGTGATGTCGGCTCA 4254
Qy 1020 ATCTTATCTGTTCAATAGTCAGTTGACAGACATTTAGGTGTTTCCACTTTTTGACCAT 1079
Db 4255 CTGCAACTCTGCTCTGGGTTCAAGCGCTGCAACCTCTGCTCTCTGGGTTCAAGCAAT 4314
Qy 1080 ATGAATAATCTCCAGTGAATTCATGATATACANTTTGTGTGGGCAATGTTTTTCATTT 1139
Db 4315 TCTGTACCTCATCCACTGAGTAGTTGGAATCACAGCGTGCAGCCCACTGCCCCAGCTA 4374
Qy 1140 TGTGGGTTTATCTAGGAGTGAATTTGCTGATCCGGGTATATTTTACAGGAGCA 1199
Db 4375 ATTTTTTTGTTATTTTAGTAGAGTGGGTTTTCACACGTTGGCCAGGCTGGTCTCGAAC 4434
Qy 1200 GTTCAGGGAGAGAAA-----ACTTGGGAAAAATGAAGCATGTTTGAATAATCA 1246
Db 4435 TCCTGGCTCAAGTATTCCTCGCTTGGCTCCCAAGTGTGGATACAGGGGTGA 4494
Qy 1247 GCAAGATGACAGGGTTTTTCGGAGTTTTATTTATATTCGTTTGA CA AATGTGCAGTTT 1306
Db 4495 GCCACCATGCCAGCTGTTTTTATTTTATTTTATTTTAAAG--GCTGGGTATGTTAGCT 4552
Qy 1307 GATGAAGATCAGGTTATCTACTAGTGAGAGTGAAGTAAAGCTGGAATAGGCGGTTCA 1366
Db 4553 CATGCCGTATCTTGAACCTTGAGAGCC--CGAGGCGAGGAGGATGCTGAGACTAGGA 4611
Qy 1367 GAGTAAATCATGAAGCACTTTGAAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATA 1426
Db 4612 GTTCAAAACCAACCTGGCCACATAGCAGGTTCTTTTAAATAATAATAATAATAAT 4671
Qy 1427 ATAAAAATCAAAATTTTTTTTTTTTTTTTGGAGAAAGAGTCTGCTCTTTCACCTGGCT 1486
Db 4672 TTTATCTTATTTATTTATTTATTTATTTATTTTGGAGACAGAGTCTGTGCGCCAGGCT 4731
Qy 1487 GGAGGCGAGTGTGTATCTCAGCTCATGCAACTTTCGCTCCCGGTTCAAGCAATTC 1546
Db 4732 GGAGTGAGTGGCGGATCTCAGTCTCATGCTCACTGCAAGCTCCGCTCTCGGTTTCAAGCAATTC 4791
Qy 1547 TCCTGCTTCAAGCTCCCAAGTAGCTGGGACTACAGGCACTTTCCACCATGCCCCAGCTGA- 1605
Db 4792 TCCTGCTTCAAGCTCCCAAGTAGCTGGGACTACAGGTCCTGCCACCATGCTGGCTAAT 4851
Qy 1606 ----TTTTGTATTTTAGTAGAGTGGGATTTCACTTTGTTGGGCAAGCTGTCTCAAACT 1663
Db 4852 TTTTTTTGTTATTTTAGTAGAGACAGGTTTCAACCGTGTAAACAGGATGTTCTCAATCT 4911
Qy 1664 TTTTGTGTCTAATTTGTTGTAATTTGTTCTTTTGTCTGAGTAGGCCCCCAGACCA 1723
Db 4912 CCTGAC-----TTCTGTATCCACCCACTCAGCCTCCCAAAAGTGTGGGATTACAGGCGT 4966

Qy 1724 AAAAAATTAATCTTAGAATCCAAATCCAAATCAGTGTGTGTGTTTGACCACTGTCACTTGAGAAC 1783
Db 4967 GAGCCACACGCTGGCCCT-----GGCTATCTCTTTTAAAACTTTATTTTGGAGAA 5019
Qy 1784 CACAGTGTGACAGGCGCTCAGGAGTAGAGTGTATCTCTCTGAAAGAGAGAAATAGATG 1843
Db 5020 AAAAAATCAGAAGTGGCATTTTGGCTTTTACATGTACAGCAATAAGTTGAAAAAAATTTT 5079
Qy 1844 AAAATATTTCTCCGGGCGCAGGCTGGTGGCTCATGCTGTAAATCCAGCACTTTTGGGAGGC 1903
Db 5080 TTTTACTGGGTGGCTGGGCGGCTCAGCCCTGTATCCAGCACTTTTGGGAGGC 5139
Qy 1904 CAAGGCAATGTGAATCA-----CCTGAGGTGAGAGT 1934
Db 5140 TGAGGCTGTGGATCATAGATCAGGAGGCTGAGGAGGCTGAGATCACAAAGTCAGGAGA 5199
Qy 1935 TCAAAACCAAGCTGGCAACATGCTGAAACCCGCTCTACTATAAATAACAAAAATTAG 1994
Db 5200 TCAGAGCAATCTCTGGCTAACGTGTGTGAACCCCATCTCTACTATAAATAC-AAAAATTAG 5258
Qy 1995 CTAAGTGTGTGGCGCATGCTGTAAATCCAGCTACTTTGGGAGGCTGAGGAGGAGAAAT 2054
Db 5259 CTGGGCTGTGTGTGATGCTGTATCCAGCTACTTTGGAAGCTGAGGAGGAGAAAT 5318
Qy 2055 TCTTGAACCCGGAGGAGGAGGTTGCAAGTGAAGCGAGATCACACCACTGCACTCCAGCCT 2114
Db 5319 GCTTGACCCAGGAGGAGGAGGTTGCAAGTGAAGCGGATATTGAGCCACTGCAATGCCAGCCT 5378
Qy 2115 GGGGAGAGAGGAGGAGGTTCTCTCAAAAAACAAAAA 2153
Db 5379 -GGCAACAGAGCAAGACTCTGTCTCAAAAAACAAAAA 5416

RESULT 10
AAL05122
ID AAL05122 standard; DNA; 13216 BP.
XX
AC AAL05122;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 7810.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
XX
OS cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN W020015320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.

PR	14-AUG-2000;	2000US-0224518P.	PR	08-NOV-2000;	2000US-0246477P.
PR	14-AUG-2000;	2000US-0224519P.	PR	08-NOV-2000;	2000US-0246478P.
PR	14-AUG-2000;	2000US-0225213P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	2000US-0225214P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246610P.
PR	18-AUG-2000;	2000US-0228273P.	PR	08-NOV-2000;	2000US-0246611P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-0226686P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249208P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249209P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249210P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249214P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249215P.
PR	06-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249216P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249217P.
PR	08-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249219P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249220P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249221P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249222P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249223P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249224P.
PR	08-SEP-2000;	2000US-0233081P.	PR	17-NOV-2000;	2000US-0249225P.
PR	12-SEP-2000;	2000US-0231968P.	PR	17-NOV-2000;	2000US-0249226P.
PR	14-SEP-2000;	2000US-0232397P.	PR	17-NOV-2000;	2000US-0249227P.
PR	14-SEP-2000;	2000US-0232398P.	PR	17-NOV-2000;	2000US-0249228P.
PR	14-SEP-2000;	2000US-0232399P.	PR	17-NOV-2000;	2000US-0249229P.
PR	14-SEP-2000;	2000US-0232401P.	PR	17-NOV-2000;	2000US-0249230P.
PR	14-SEP-2000;	2000US-0232401P.	PR	17-NOV-2000;	2000US-0249231P.
PR	14-SEP-2000;	2000US-0232401P.	PR	17-NOV-2000;	2000US-0249232P.
PR	14-SEP-2000;	2000US-0232401P.	PR	17-NOV-2000;	2000US-0249233P.
PR	14-SEP-2000;	2000US-0232401P.	PR	17-NOV-2000;	2000US-0249234P.
PR	14-SEP-2000;	2000US-0232401P.	PR	17-NOV-2000;	2000US-0249235P.
PR	14-SEP-2000;	2000US-0232401P.	PR	17-NOV-2000;	2000US-0249236P.
PR	14-SEP-2000;	2000US-0232401P.	PR	17-NOV-2000;	2000US-0249237P.
PR	14-SEP-2000;	2000US-0232401P.	PR	17-NOV-2000;	2000US-0249238P.
PR	14-SEP-2000;	2000US-0232401P.	PR	17-NOV-2000;	2000US-0249239P.
PR	14-SEP-2000;	2000US-0232401P.	PR	17-NOV-2000;	2000US-0249240P.

1501 TGATCTCAGTCTCACTCAACTTTTCGCTCCCGGTTCAAGCAATTTCTCTGCTTCAGCCT 1560
1561 CCACAGTAGCTGGAGTACAGGCACTTCCACCATCCAGCTGATTTTGTGATTTTATAG 1620
1621 TAGAGATGGATTTCACTTTGTTGGCAAGCTGGTCTCAAACTTTTGTGTCATATTTG 1680
1680 TACAGATGGGTTTCCACATGTTGGTCAAGCTGGTCTCGAATCTCCGACTTCATGATCCG 3864
1691 TTGTAATATTGTTCTTTTCTGCTGAGCTAGGCGCCCGACACCAAAAAAATAAATCTTAG 1740
1741 AATCCAAATCAGTGTGTGTTGACCATGTCACTTGAGAACCAAGGATGATGATGATGATG 1800
1801 CTGAGAGTAGAGTATCTCTCTCGAAGAGAAATAGAAATATTTCTCCGGGCC 1860
1861 AGCGTGTGGCTCATGCTGTATATCCAGCACTTTGGGAGGCCAAGGATGATGATGATGAT 1920
1921 CTGAGTTCAGAGTTCAAAACCAAGCTGGCCCAATGATGATGATGATGATGATGATGAT 1980
1981 ATACAAAATTAAGTAAAGTGTGGTGGCGCATGCTGATATCCAGTCTTCTGAGGGGT 2040
2041 GAGGAGGAGAAATTTCTTGAACCGGAGGAGGAGGTTGAGTGAAGCGAGATCACACCA 2100
2101 CTGCACTCCAGCTGGGGGAGAGAGAGAGATCTTCTCTCAAAAAAACAACAAAGAA 2160
2161 ATTAAGCAATTAAGCA 2177
4334 ACAAACAACAACAACA 4350

RESULT 11

ABL98014 ID ABL98014 standard; DNA; 13216 BP.

XX ABL98014;

XX AC

XX 21-JUN-2002 (first entry)

DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2666.

XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.

XX Homo sapiens.
OS

XX WO200155317-A2.
PN

XX 02-AUG-2001.
PD

XX 17-JAN-2001; 2001WO-US001329.
XX PF

XX 31-JAN-2000; 2000US-0179065P.
XX PR

PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228242P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232401P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.

RESULT 12
ABV52505/c
ID ABV52505 standard; cDNA; 396 BP.
XX
AC ABV52505;

DK
XX
XX
PT
PT
XX
PS

XX DT 17-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 52496.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 10183; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 396 BP; 125 A; 96 C; 57 G; 118 T; 0 U; 0 Other;

Query Match 8.8%; Score 348.6; DB 5; Length 396;

Best Local Similarity 98.9%; Pred. No. 5.1e-62;

Matches 351; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1184 TATTTGACAGGAGTTCAGGGGAGAAACCTTGGGAAATGAAGCATGTTAGAAA 1243

Db TTTTTCAGGAGAGTTCAGGGGAGAAACCTTGGGAAATGAAGCATGTTAGAAA 329

Qy 1244 TCAGCAGAGTGCAGGGGTTTTCGGAGTTTATTTATATCTGTTGACAAATGTCGAG 1303

Db TCAGCAGAGTGCAGGGGTTTTCGGAGTTTATTTATATCTGTTGACAAATGTCGAG 269

Qy 1304 TTTGATGAAGATACAGGTATACCTAAGTGAGAGTGAATTAAGGCTGGAATAGGGCGT 1363

Db TTTGATGAAGATACAGGTATACCTAAGTGAGAGTGAATTAAGGCTGGAATAGGGCGT 209

Qy 1364 TCAGAGTAAATCATGAGACCTTTGATACCAAAATTAAGGAGCTTGGCTGTAAACAAA 1423

Db TCAGAGTAAATCATGAGACCTTTGATACCAAAATTAAGGAGCTTGGCTGTAAACAAA 149

Qy 1424 ATAATAAAAAATCAAAATTTTTTTTTTTTTTTTTTTTGTAGAAAGTCTGCTCTTCCACCTG 1483

Db ATAATAAAAAATCAAAATTTTTTTTTTTTTTTTTTTTGTAGAAAGTCTGCTCTTCCACCTG 89

Qy 1484 GCTGGAGGCGAGTGTGATCTCAGCTCACTGCAACTTTTCGCTCCCGGTTCA 1538

Db 88 GCTGGAGGCGAGTGTGATCTCAGCTCACTGCAACTTTTCGCTCCCGGTTAA 34

RESULT 13

AA05121

ID AAL05121 standard; DNA; 29163 BP.

XX AAL05121;

XX 21-NOV-2001 (first entry)

XX Human reproductive system related antigen DNA SEQ ID NO: 7809.

XX Human; reproductive system related antigen; reproductive system disorder;

XX cancer; gene therapy; ds.

XX Homo sapiens.

XX WO200155320-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001339.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180282P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

XX 14-AUG-2000; 2000US-0225759P.

XX 18-AUG-2000; 2000US-0226279P.

XX 22-AUG-2000; 2000US-0226681P.

XX 22-AUG-2000; 2000US-0226868P.

XX 23-AUG-2000; 2000US-0227182P.

XX 23-AUG-2000; 2000US-0227009P.

XX 30-AUG-2000; 2000US-0228924P.

XX 01-SEP-2000; 2000US-0229287P.

XX 01-SEP-2000; 2000US-0229343P.

XX 01-SEP-2000; 2000US-0229344P.

XX 01-SEP-2000; 2000US-0229345P.

XX 05-SEP-2000; 2000US-0229509P.

XX 05-SEP-2000; 2000US-0229513P.

XX 06-SEP-2000; 2000US-0230437P.

XX 06-SEP-2000; 2000US-0230438P.

XX 08-SEP-2000; 2000US-0231242P.

XX 08-SEP-2000; 2000US-0231243P.


```
QY 1921 CTGAGGTCAGAGTTCAGAACCCAGCCTGGCCACATGGTGAAACCCCGTCTCTACTAAAA 1980
Db 4093 CTGAGGTCGCGAGTTCAGAACCCAGCCTGGCCACATGGTGAAACCCCGTCTCTACTAAAA 4152
QY 1981 ATACAAAAATTAGCTAAGTGTGTGGCGCATGCTCTAATCCAGCTACTTGGGAGGGT 2040
Db 4153 ATAC-AAAAATAGCCACGCTGTGTGGCGCATGCTCTAATCCAGCTACTTGGAGGCT 4211
QY 2041 GAGCAGGAGAAATTTCTGAACCCGCGGAGGAGAGGTTGAGTGAGCCGAGATCACACCA 2100
Db 4212 GAGCAGGAGAAATCACTTGAACCCAGGAGGCGGAGTTGAGTGAGCCGAGATCACACCA 4271
QY 2101 CTGCACCTCCAGCCTGGGGGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAGA 2160
Db 4272 CTGCACCTCCAGCCTGGGGGAGAGCGAGTGAGGCTGTGTCTCAAAAAACAAAAACAAA 4331
QY 2161 ATTAAGCAAAATTAGACA 2177
Db 4332 ACAAAACAAACAAACA 4348

RESULT 14
ABL98013
ID ABL98013 standard; DNA; 29163 BP.
XX
AC ABL98013;
XX
21-JUN-2002 (first entry)
XX
Human testicular antigen encoding DNA fragment SEQ ID NO: 2665.
DE
XX
Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200155317-A2.
XX
02-AUG-2001.
XX
17-JAN-2001; 2001WO-US001329.
XX
31-JUN-2000; 2000US-0179065P.
PR
04-FEB-2000; 2000US-0180628P.
PR
24-FEB-2000; 2000US-0184664P.
PR
02-MAR-2000; 2000US-0186350P.
PR
16-MAR-2000; 2000US-0189874P.
PR
17-MAR-2000; 2000US-0190076P.
PR
18-APR-2000; 2000US-0198123P.
PR
19-MAY-2000; 2000US-0205515P.
PR
07-JUN-2000; 2000US-0209467P.
PR
28-JUN-2000; 2000US-0214886P.
PR
30-JUN-2000; 2000US-0215135P.
PR
07-JUL-2000; 2000US-0216647P.
PR
07-JUL-2000; 2000US-0216880P.
PR
11-JUL-2000; 2000US-0217487P.
PR
11-JUL-2000; 2000US-0217496P.
PR
14-JUL-2000; 2000US-0218290P.
PR
26-JUL-2000; 2000US-0220963P.
PR
26-JUL-2000; 2000US-0220964P.
PR
14-AUG-2000; 2000US-0224518P.
PR
14-AUG-2000; 2000US-0224519P.
PR
14-AUG-2000; 2000US-0225213P.
PR
14-AUG-2000; 2000US-0225214P.
PR
14-AUG-2000; 2000US-0225266P.
PR
14-AUG-2000; 2000US-0225267P.
PR
14-AUG-2000; 2000US-0225268P.
PR
14-AUG-2000; 2000US-0225270P.
PR
14-AUG-2000; 2000US-0225447P.
PR
14-AUG-2000; 2000US-0225757P.
PR
14-AUG-2000; 2000US-0225758P.
PR
14-AUG-2000; 2000US-0225759P.
PR
14-AUG-2000; 2000US-0225759P.
PR
18-AUG-2000; 2000US-0226279P.
PR
22-AUG-2000; 2000US-0226681P.
PR
22-AUG-2000; 2000US-0226868P.
PR
22-AUG-2000; 2000US-0227182P.
PR
23-AUG-2000; 2000US-0227009P.
PR
30-AUG-2000; 2000US-0228924P.
PR
01-SEP-2000; 2000US-0229287P.
PR
01-SEP-2000; 2000US-0229343P.
PR
01-SEP-2000; 2000US-0229344P.
PR
01-SEP-2000; 2000US-0229345P.
PR
05-SEP-2000; 2000US-0229509P.
PR
06-SEP-2000; 2000US-0229513P.
PR
06-SEP-2000; 2000US-0230437P.
PR
06-SEP-2000; 2000US-0230438P.
PR
08-SEP-2000; 2000US-0231142P.
PR
08-SEP-2000; 2000US-0231143P.
PR
08-SEP-2000; 2000US-0231244P.
PR
08-SEP-2000; 2000US-0231413P.
PR
08-SEP-2000; 2000US-0231414P.
PR
08-SEP-2000; 2000US-0232080P.
PR
08-SEP-2000; 2000US-0232081P.
PR
12-SEP-2000; 2000US-0231968P.
PR
14-SEP-2000; 2000US-0232397P.
PR
14-SEP-2000; 2000US-0232398P.
PR
14-SEP-2000; 2000US-0232399P.
PR
14-SEP-2000; 2000US-0232400P.
PR
14-SEP-2000; 2000US-0232401P.
PR
14-SEP-2000; 2000US-0233063P.
PR
14-SEP-2000; 2000US-0233064P.
PR
14-SEP-2000; 2000US-0233065P.
PR
21-SEP-2000; 2000US-0234223P.
PR
21-SEP-2000; 2000US-0234274P.
PR
25-SEP-2000; 2000US-0234997P.
PR
25-SEP-2000; 2000US-0234998P.
PR
26-SEP-2000; 2000US-0235484P.
PR
27-SEP-2000; 2000US-0235834P.
PR
27-SEP-2000; 2000US-0235836P.
PR
29-SEP-2000; 2000US-0236327P.
PR
29-SEP-2000; 2000US-0236367P.
PR
29-SEP-2000; 2000US-0236368P.
PR
29-SEP-2000; 2000US-0236369P.
PR
02-OCT-2000; 2000US-0236802P.
PR
02-OCT-2000; 2000US-0237037P.
PR
02-OCT-2000; 2000US-0237038P.
PR
02-OCT-2000; 2000US-0237039P.
PR
13-OCT-2000; 2000US-0239935P.
PR
13-OCT-2000; 2000US-0239937P.
PR
20-OCT-2000; 2000US-0240960P.
PR
20-OCT-2000; 2000US-0241221P.
PR
20-OCT-2000; 2000US-0241785P.
PR
20-OCT-2000; 2000US-0241786P.
PR
20-OCT-2000; 2000US-0241787P.
PR
20-OCT-2000; 2000US-0241808P.
PR
20-OCT-2000; 2000US-0241809P.
PR
20-OCT-2000; 2000US-0241826P.
PR
01-NOV-2000; 2000US-0244617P.
PR
08-NOV-2000; 2000US-0246474P.
PR
08-NOV-2000; 2000US-0246475P.
PR
08-NOV-2000; 2000US-0246476P.
PR
08-NOV-2000; 2000US-0246477P.
PR
08-NOV-2000; 2000US-0246523P.
PR
08-NOV-2000; 2000US-0246524P.
PR
08-NOV-2000; 2000US-0246525P.
PR
08-NOV-2000; 2000US-0246526P.
PR
08-NOV-2000; 2000US-0246527P.
PR
08-NOV-2000; 2000US-0246528P.
PR
08-NOV-2000; 2000US-0246532P.
PR
08-NOV-2000; 2000US-0246609P.
PR
08-NOV-2000; 2000US-0246610P.
```


PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234283P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249266P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249399P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465566/50.

Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases.

Disclosure; SEQ ID NO 2140; 1180pp; English.

The present invention relates to the isolation of novel human enzyme polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the

Db	2364	TAGCCATTTTACAGATAGGACCAAGTAGGAGNCTTAGGGAAT-----	2323
Qy	1203	CAGGGGAGAGAAAACCTTGGGAAAAATGAAGCATGTTTAGAAATCAGCAAG-AGTCAGGGG	1261
Db	2322	-----TAGAGATTTAGGGAATCTGCTAACGGTTGCACAGCCAGGAATATAGGCAGCT	2269
Qy	1262	TTTTTCGGAGTTTTATTTATATCTCTGTGACAAATGTGCAAGTTGATGAGATACAAGT	1321
Db	2268	AGAAAAGCTTTCAGCTAGGCATGGAGCTCACACTGTAATCTCAGTGCCTTTGGGAGG	2209
Qy	1322	TATACTAAGTGAGAACTGAGAATTAGGCTGGNAATAGGGCGTTTCAGAGTAAATCATGAA	1381

1382	QY	GCATTTTGAATACCAAAAATTAGGAGCTTGGCTGTATAA	1441	CAAAATATAAAAAATCACAAT
2148	Db	ACCTATCTCTATGAAAAATAAACAAATAAAAAAAGTGTAGAAATATCTTGCTTTTCTT	2089	
1442	QY	TTTTTTTTTTTTTTTTCAGAAAAAGTCTTGCTCTTTTACCTGGCTGGAGGCGAGTGGTGT	1501	
2088	Db	TTCTTTTTTTTTTTTTTTTGGACAGAACTCTACTCTGTGTTGCCAGGCTGAGTGAATGGTGT	2029	
1502	QY	GATCTCAGTCTCAGTCAACTTTTCGCTCCCGGGTTCAAGCAATCTCTGCTTCAGCCTC	1561	
2028	Db	AATCTCAGCTCACCAAAACCTCGCCTCTCAGGTTTCAAGTGATCTCTCCCTCCCTCAGCCTC	1569	
1562	QY	CCAAAGTAGCTGGAGCTACAGGCATTTCCCAACCATGCCAGTCAATTTTGTATTTTGTAGT	1621	
1968	Db	CTGAGTAGTTGGAGCTTACAGGCGGGTGCACCATGTCCGGCTAATTTTGTATTTTGTAGT	1909	
1622	QY	AGAGATGGGAAATTCATTTTGTGGCCAAAGCTGTCTCAAACTTTT	1666	
1908	Db	AGAGACGGGGTTTCACTATGTTGACCAAGGTGCTCTCGAACTCCTGACCTTGATCTGC	1849	
1667	QY	-----TGCTGTCTAATTTCTGTAA-----CTATTG	1692	
1848	Db	CCGCTCGGCTCCCAAGTGTGGGATACAGTGTAGCCAGCTCGCTGCGCGCTCTT	1789	
1693	QY	TTCTTTTGTGTAGGTAGGCCCCCAGACCAAAAAATAAATCTTAGAATCCAAATCAG	1752	
1788	Db	GGTTTTTCTTTACAAACACGCTGGCCAGCCAGATATCTTTATTTCCGCGAAAGAAACCCCTC	1729	
1753	QY	TGTGTTGGTTTG--ACCACTGTCTACTTTGAGAACACAGTGTGACAGGCGCTCAGGAGT--	1809	
1728	Db	GGTGGAGATAGCAAGACAGTTACTCAAGATCATTTGATATAACCGGCTAGTTGTAAATTC	1669	
1810	QY	--AGAGTGATCTCTGCTCCAAAAGAGAAATPAGAAATATAATTTCTCCGGCCAGGCGTG	1867	
1668	Db	CAACATTTGATCTGACCTCCAGGATGCACCTGTTAGAAAGTGGATTTGGCCAGGCGTG	1609	
1868	QY	--GTGGCTCATGCTGTAAATCCAGACATTTTGGAGGCCAAGGCATGTGCATCACCTGAG	1925	
1508	Db	CTGTGGCTCACCTGTAACTCTAGCGCTTTGGAGGCAGAGGCGCTGATTTGCTTGAG	1549	
1926	QY	GTCAAGGATTTCAAAACCAAGCCTGGCCAAACATGTGTAAACCCCGCTCTCTACTAAAAATACA	1985	
1548	Db	GTCAGTAGTTAATAATAGCTGGTCAAATGTTGAAACCTCATCTACTAAAAATACA	1489	
1986	QY	AAAAATTAGTAAATGTGGTGGGCCATGTCCTGTAAATCCAGCTACTTTGGAGGGGTGAGGC	2045	
1488	Db	AAATTAGCC--AGGCGTGTGGCACCCACCTATAGTCCCAAGTACTTTGGGGGGCTGAGGC	1431	
2046	QY	AGGAGAAATTTTCMAACCCGGAGGCGAGAGTTTGCAGTGAAGCGAGATCACACCTGCA	2105	
1430	Db	AGGAGAACTGCTTAGCCCGGGAGGTGGAGTTTGCAGGAGCCAGATC--TGCCACTGCA	1372	
2106	QY	CTCCAGCCTGGGGAGAGACGAGCTTCTCTCAAAAAACAAAAACAAAAAGAT	2162	
1371	Db	CTCTAGCCTGAGGGACAGACGAGCTCCATCTCAAAAAAGAAAAAGAAAAAGT	1315	

Search completed: May 12, 2004, 19:47:11

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:19 ; Search time 284.726 Seconds
(without alignments)
7714.406 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3958

Perfect score: 3958

Sequence: 1 aggaattcattccatttaaat.....gggacctccacgcacgcgg 3958

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349.8	8.8	12394	4	US-09-488-856A-10
2	345	8.7	84495	4	US-09-797-906-3
3	322.4	8.1	3805	4	US-09-108-006C-3
4	321	8.1	35060	3	US-08-814-095-7
5	319.8	8.1	392000	4	US-10-027-983-11
6	316.6	8.0	59065	4	US-09-813-817-3
7	316.6	8.0	59065	4	US-09-978-157-3
8	312.4	7.9	174493	4	US-08-804-471A-3
9	312.4	7.9	174493	4	US-10-238-709-3
10	308.2	7.8	162450	4	US-09-345-882-1
11	307.6	7.8	43950	4	US-09-735-934A-3
12	307.6	7.8	43950	4	US-10-060-332-3
13	307.2	7.8	70000	4	US-09-851-896-3
14	304	7.7	39982	4	US-08-820-924-3
15	303	7.7	43950	4	US-09-735-934A-3
16	303	7.7	43950	4	US-10-060-332-3
17	301.8	7.6	246240	2	US-08-724-394A-20
18	301.8	7.6	246240	2	US-08-724-394A-21
19	301.8	7.6	246240	2	US-08-724-394A-22
20	299.8	7.6	39982	4	US-09-820-924-3
21	297.2	7.5	70000	4	US-09-851-896-3
22	296.6	7.5	59065	4	US-09-813-817-3
23	296.6	7.5	59065	4	US-09-978-157-3
24	291.2	7.4	83450	4	US-09-811-469-3
25	289.2	7.3	55298	4	US-09-491-356C-1
26	288	7.3	21234	4	US-09-810-671-3
27	288	7.3	21234	4	US-10-109-854-3

28	287.4	7.3	116592	4	US-09-818-512-3	Sequence 3, Appli
29	286.8	7.2	35060	3	US-08-814-095-7	Sequence 7, Appli
30	285.8	7.2	162450	4	US-09-345-882-1	Sequence 1, Appli
31	283.8	7.2	12394	4	US-09-488-856A-10	Sequence 10, Appli
32	282.8	7.1	3805	4	US-09-108-006C-3	Sequence 3, Appli
33	282.8	7.1	246240	2	US-08-724-394A-20	Sequence 20, Appli
34	282.8	7.1	246240	2	US-08-724-394A-21	Sequence 21, Appli
35	282.8	7.1	246240	2	US-08-724-394A-22	Sequence 22, Appli
36	280	7.1	116592	4	US-09-818-512-3	Sequence 3, Appli
37	275	6.9	83450	4	US-09-811-469-3	Sequence 3, Appli
38	275	6.9	84495	4	US-09-797-906-3	Sequence 3, Appli
39	270.4	6.8	72604	4	US-09-268-992-7	Sequence 7, Appli
40	270.4	6.8	72604	4	US-09-657-474-7	Sequence 7, Appli
41	269.4	6.8	23187	4	US-09-499-522-1	Sequence 1, Appli
42	267.2	6.8	45546	4	US-09-146-053-6	Sequence 6, Appli
43	265.6	6.7	21721	4	US-09-269-939A-41	Sequence 41, Appli
44	264	6.7	44848	4	US-09-435-739-42	Sequence 42, Appli
45	263	6.6	98844	4	US-09-791-211-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1
US-09-488-856A-10
; Sequence 10, Application US/09488856A
; Patent No. 6316259
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Robert McKay
; APPLICANT: Madeline M. Butler
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF GLYCOGEN SYNTHASE KINASE 3 ALPHA EXP
; FILE REFERENCE: RTS-0115
; CURRENT APPLICATION NUMBER: US/09/488,856A
; CURRENT FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 10
; LENGTH: 12394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(397)
; NAME/KEY: CDS
; LOCATION: (2438)...(2625)
; NAME/KEY: CDS
; LOCATION: (5639)...(5722)
; NAME/KEY: CDS
; LOCATION: (5864)...(5974)
; NAME/KEY: CDS
; LOCATION: (7902)...(8032)
; NAME/KEY: CDS
; LOCATION: (8121)...(8227)
; NAME/KEY: CDS
; LOCATION: (9197)...(9294)
; NAME/KEY: CDS
; LOCATION: (9375)...(9470)
; NAME/KEY: CDS
; LOCATION: (9898)...(10084)
; NAME/KEY: CDS
; LOCATION: (10431)...(10523)
; NAME/KEY: CDS
; LOCATION: (11713)...(11786)
US-09-488-856A-10

Query Match 8.8%; Score 349.8; DB 4; Length 12394;

Best Local Similarity 55.2%; Pred.No. 4.9e-77;

Matches 1016; Conservative 0; Mismatches 742; Indels 81; Gaps 14;

Qy 370 TATTATTAGACAGGTTCTCCTCTGTCCACCCAGGCTGGAGTGCAGTGGCACAATCATATA 429

Db 3604 TTTTITGAGATGGAGTCTTACTCTGTCCACCCAGGCTGGAGTGCAGTGGCACAATCTTG 3663

Query Match 8.7%; Score 345; DB 4; Length 84495;
Best Local Similarity 55.0%; Pred. No. 2.1e-75;
Matches 9/9; Conservative 0; Mismatches 715; Indels 85; Gaps 12;

445 AAACCTCTGGGCTCAAGTATCTTACTACCTCAGCC-TCAGAGTAGCTAGGACTACAGG 503
18461 AACTTCTGGGCTCAAGCCATCCACCTGCCTGGCTTCCAAAGTGGTGGGATTAACAAG 18402
504 CACACAGGCAATACCTGGCTAAATTTTTTTTAAATTTTCAATTTATGATTTCAATTTTC 563
18401 CATTAGTCACACGCCCTGGACTATTTATATATATATATATATATATATATATATTTT 18342
564 TTCTCTTTTGG-TTGT 622
18341 TTTTCTTTTGT 18282
623 GCAGTGGCATGGTGACAGCTGA--GCAGCCCTTGACTTCTGGGCTCAAGTGATCTCTCTG 680
18281 GCAGTGGCATGATCTCAGCTCAGCTGCAACCTTGAGCTCTCTGGGTTCAAGCAATTTCTCTG 18222
681 CCTCAGCTCCCAAGTAGCTGGGACTCAACACAGCTGTACCATGTCTGGCTGGCTGATTTTT 740
18221 CCTCAGCTCTTAAAGTAGCTGGGATTAAGGCAATGTACCGGCGCCAGCTAAATTTTT 18162
741 TTTTCTTGAACAGGGTATCACTCTGTGTGCCAGGCTGGAGTACAGTGGGCTTAATAATAG 800
18161 TTTTCTT-----TTTTTGT 18109
801 CTCACCTCAGCTCCCTCTGGGCTCAAGCAATCCGCTGGCTCAGCATCTCTGAGTAGC 860
18108 CCAGGCGCATCTCGAATCTCTGACCTCAAGTATCTGCCGCTCAGCTCCCAAGTGT 18049
861 TGGGACTACAGGCTGT 920
18048 TAAGATTTACAGGATGAGCACCACCTGCGGCTCTCTATTTTATAGTTTAAAGACACA 17989
921 GAGGAGCTGT 980
17988 GA-----GTCCCATCTGT 17933
981 GTTGCCATGATCCGCCCATCTCCACTTCCCAAGTGTCTATCTTATCTG--TTCAATAGT 1038
17932 GCCTTAACTCTCTGGCTAAAGCAATCTCTACCTCAGCTCCGAGTAGCTACATGCAC 17873
1039 CAGTTGACAGCATTTAGT 1098
17872 ATGCCACCATACCCAGCTAAATTTTAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 17813
1099 ATATTCATGTATACATTTGTGTGGGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1158
17812 CCAGGCTGTGTGGAACTCTGTGGCTTAAGCAATCTCCCATCTTGGCTTCCAAAGTG 17753
1159 AGTGGAAATGCTGGATCCCGGTAAATTTTGAAGGAGAGTTCAGGCGAGGTTTCAAGGAA 1218
17752 CTGGGATTAAGGTGTGAGCCACCATGTCTGGCCACCAATGTACTTTTCAAGGTCACTT 17693
1219 TGGGAAATGAGCATGT 1278
17692 GACTGCACTGTGG--AGAATAAATCGAGGAGGATTAATAGTTTAAATAGGCGAGTATTTA 17635
1279 TTATATCTGTGTGCAAAATGTGCAAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1338
17634 ATGACCATACAGAGCCATCTCACTTTAAGTAAATCATTAATTTAATCTGATCATTTCT 17575
1339 GAGAAATTAAGGCTGGAATFAGGCGCTTCAGAGTAAATCATGAGCACTTTGA-----ATAC 1394
17574 CTATTTTATGGACATTTAGACTATTTCCAGTCTTTTGTGTGTGTGTGTGTGTGTGTGTGT 17515
1395 CAAAATTAAGAGCTTGT 1454
17514 TGTAAATAACATTTCTGAAACCTAAATCTTTGGGAGCATCTCTAATTAATTTCTTCTCTT 17455
1455 TTGAGAAAGAGTCTTGTCTTTTCCCTCTGCTGGCTGGAGGCGAGTGTGTGTGTGTGTGTGT 1514

17454 TTTTGAACAGAGCTCCTACCTGTTGCCAGGCTCGAGTGCAGTGTGTGTGTGTGTGTGTGT 17395
1515 TGAACCTTTGGCTCCCGGTTCAAGCAATCTCTGTCTTCCAGCTTCCAGCTTCCAGTGTGGG 1574
17394 TGAACCTTCCACCTCTCTGGGTTCAAGTGTATCTCTGTCTTCCAGCTTCCAGTGTGGG 17335
1575 ACTACAGGCACTTCCACCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1634
17334 ATTACAGGTTGGGCGCCACCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 17275
1635 C-ACTTTGTGGCAAGCTGTCTCTCAAACTTTT----- 1666
17274 CGCAATTTTGGCAGGCTGTCTCAAAATCTCTGACCTCTGGGATTACAGCGGTGAGCCA 17215
1667 -----TGCTGTCAATATGTTGTAACTATTTGTCTTCTTGTCTGTGTGTGTGTGTGTGT 1718
17214 CCAGCGCCGACTTAATTTTCTAAATATAGATTTCTAGAGTGGAAATTAAGTAATAA 17155
1719 GACCAAAAATAAATAAATCTTAGAATCCAAATCAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1778
17154 GTTAAAGGTTATGAATTTTGTGTGTATTTTAAATAATTTTAAAGATTAAGATTT 17095
1779 AGAACACAGTGTGACAGGCTCCTCAGGAGTAGAGTGTCTCTGTCTGAAAGAGAAATA 1838
17094 TTGATCATTAATAATAAATAATCTACTTTGCAAAATGGAATCTTTTCACTATCCAGTC 17035
1839 GAAT-----GAAATATTTCTCCGGCCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1874
17034 AAATTTATCAGCGGAAATGTCTTTTAAATAAGATGAAAACTAGGCCAGCGAGTGGCTG 16975
1875 ATGCTGTATATCCAGCACTTTTGGGAGGCAAGCATGTGTGTGTGTGTGTGTGTGTGTGTGT 1934
16974 ACCTGTGTATCCAGCACTTTAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 16917
1935 TCAAAACAGCTGTGGCAACATGT 1994
16916 TCAGAGCACTCTGGCTAAACATGT 16857
1995 CTAGT 2054
16856 CCGGT 16797
2055 TCTTGAACCCGGAGGCGAGAGT 2114
16796 TCTTGAACCCGGAGAGT 16737
2115 GGGGAGAGAGCGAGAGT 2153
16736 GGGCAACAGGCGCAGACTCTGT 16698

RESULT 3

US-09-108-006C-3/c
; Sequence 3, Application US/09108006C
; Patent No. 6524613
; GENERAL INFORMATION:
; APPLICANT: Steer, Clifford J.
; Kren, Betsy T.
; Bandyopadhyay, Paramita
; Roy-Chowdhury, Jayanta
; TITLE OF INVENTION: Hepatocellular Chimeraplasty
; ADDRESS: Kimeragen, Inc.
; STREET: 300 Pheasant Run
; CITY: Newtown
; STATE: PA
; COUNTRY: USA
; ZIP: 18940
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible


```
; NAME/KEY: promoter
; LOCATION: 4089..22464
; OTHER INFORMATION: /function= "ACHE Promotor"
; OTHER INFORMATION: /standard_name= "ACHE Promotor"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 22465..22537
; OTHER INFORMATION: /function= "non-translated"
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 24090..25177
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "(translation start:"
; OTHER INFORMATION: 24110)"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25524..26009
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27005..27274
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27255..28007
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 5
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 27385..27387
; FEATURE:
; NAME/KEY: exon
; LOCATION: 28008..28129
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 6
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 28129..28131
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (34528..34895)
; OTHER INFORMATION: /function= "arsenite resistance
; OTHER INFORMATION: gene"
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (34092..34358)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33779..33963)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33493..33591)
```

```
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33297..33408)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 5
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32959..33094)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 6
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32569..32628)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 7
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32386..32468)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 8
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31894..32080)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 9
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31363..31534)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 10
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31131..31284)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 11
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30816..31011)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 12
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30470..30626)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 13
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30187..30274)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 14
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29945..30073)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 15
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29664..29856)
; OTHER INFORMATION: /gene= "ARS"
; OTHER INFORMATION: /number= 16
; US-08-814-095-7
```

Query Match 8.1%; Score 321; DB 3; Length 35060;
Best Local Similarity 68.8%; Pred. No. 1.3e-69;
Matches 511; Conservative 0; Mismatches 205; Indels 27; Gaps 4;

```
QY 1445 TTTTCTTTTTCAGAAAGAGTCTTCTTTTCACTGCTGAGGGGAGTGGTGTGAT 1504
      |||||
Db 9802 TCTTTTCTTTTTCAGAGCGGAGTCTCTATCGCCAGGCTGGAGTGCAGTGCACAAT 9861
      |||||
QY 1505 CTCAGCTCACTGCACATTTTCGCTCCCGGTTTCAAGCAATTCTCTCTTCAGCTCCCA 1564
```



```
Db 11689 CAGCCTTGGGGAACAGAGTGAGACTCTGTCTAAACAAAAACAAACAAAAATC 11630
QY 2168 A 2168
Db 11629 A 11629

RESULT 7
US-09-78-197-3/c
; Sequence 3, Application US/09978197
; Patent No. 6403353
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001178DIV
; CURRENT APPLICATION NUMBER: US/09/978,197
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/813,817
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-78-197-3

Query Match 8.0%; Score 316.6; DB 4; Length 59065;
Best Local Similarity 66.7%; Pred. No. 2.2e-68;
Matches 561; Conservative 0; Mismatches 264; Indels 16; Gaps 7;

QY 1339 GAGAAATTAAGGCTGGATAGGGGCTTCAGAGTAAATCATGAAGCACTTTGAATACCAA 1398
Db 12464 GAGAGTGAGGCTGAAGTGCAGGAGCAGGAGTTCTGGAGATGGAGAGAGGCCACAT 12405

QY 1399 ATTAAGGAGCTTGGCTGTAAACAAAAATAAATAAATCAAAATTTTTTTTTTTTGA 1458
Db 12404 TACGGAGGCTCTACTGTAGGCCAAATTA-AGGAATTTTTTTTTTTTTTTTTTTTGA 12346

QY 1459 GAAAGTCTTGTCTTTTCACTGTGCTGGAGGAGTGTGTGATCTCAGCTCACTGCA 1518
Db 12345 GATGGAGTCTTACTCTGTCAACA-GGCTGGAGTGCAGTGCCTGATCTGGGCTCACTGCA 12287

QY 1519 ACTTTGCGCTCCGGGTTCAAGCAATTCCTGCTTCAGCCTCCCAAGTAGCTGGACTA 1578
Db 12286 ACTCTGCTCCCGAGTTCAAGTGATCTCTGCTCAGCCTCCCAAGTAGCTGGACTA 12227

QY 1579 CAGGCATCTCCACCATGCCAGCTGAATTTTGTATTTTATTTAGTAGAGATGGGATTCAT 1638
Db 12226 CAGGCACAGCCACCAACGACCTAAATTTTGTATTTTATTTATTTAGTAGAGATGGGATTCAT 12167

QY 1639 TTGTGGCCAAAGTGTCTCAACTTTTGTGCTGCATATTT-----GTGTAACT 1688
Db 12166 ATGTTGGCCAGGATGTCTCAATCTTTGACTTCATGATCTGCTGCCTCGGCTCCCAA 12107

QY 1689 ATTTGTTCTTTTGTGAGGTAGGCGCCCGCCAGACCAAAAAATAAATCTTAGAATCCAA 1748
Db 12106 AGTGCTGAGATGATAGTGTGAGCCACCGCGCTGCGCAAGTAAAGATTTTGGTCTTTA 12047

QY 1749 TCAGTGTGTGTTTGAACACTGTCTACTTGAGAACACAGTGTGACAGGGCTCAGGAG 1808
Db 12046 TCACAAGAACTGGGAAGTTGTGAAGTATACAAGGTAGGAGAGACAACACAGAGTGACA 11987

QY 1809 TAGAGTGTATCTGTCTCGAAGAGAAATAGATGAATATATCTCCGGCGCAGGGGTGG 1868
Db 11986 CAGTCAAGTTTTTTTTTTTTTTTTTTTTTTTAAAGATCA-CTCCTTCCAGCCAGGTGCGAG 11928

QY 1869 TGGCTCATGCTGTAATCCAGCACTTTGGAGGCCCAAGGCATGTGGATCACTGAGGTTC 1928
Db 11927 TGGCTCACACCTGTAATCCAGCACTTTGGAGGCCCGAGGCGAGGTGATCACTTGAAGTTC 11868
```

```
QY 1929 AGGAGTTCAAACACAGCCTGGCCAACTGGTGAACCCCGTCTCTACTAAAAATACAAA 1988
Db 11867 AGGAGTTCGAGACTAGCTTGGCCAACTGGCCAAACCCCATCTCTACTAAAAATA-TAAA 11809

QY 1989 AATAGCTAAGTGTGCTGGCGCATCCCTGTATCCAGGTACTTGGGAGGTGAGGCAGG 2048
Db 11808 AATAGCTGGGTATGGGGCCGGTGCCT-TAATCCAGCTACTTGGGAGGTGAGGCAGG 11750

QY 2049 AGAATTTCTTGAACCCGGGAGGACAGAGTTCAGTGAAGCGAGATCACACCACTGCATC 2108
Db 11749 AGAATTTCTTAAACCCGGGAGGACAGAGTTCAGTGAAGCGGAAATGTGCCACTGCATC 11690

QY 2109 CAGCC-TGGGGAGAGAGCAGACTTCCTCTCAAAAAACAAAAACAAAGAAATTAAGC 2167
Db 11689 CAGCCTTGGGAAACAGAGTGAGACTCTGTCTAAACAAAAACAAACAAACAAAAATC 11630

QY 2168 A 2168
Db 11629 A 11629

RESULT 8
US-09-804-471A-3
; Sequence 3, Application US/09804471A
; Patent No. 6479269
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164
; CURRENT APPLICATION NUMBER: US/09/804,471A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(174493)
; OTHER INFORMATION: n = A,T,C or G
US-09-804-471A-3
```

```
Query Match 7.9%; Score 312.4; DB 4; Length 174493;
Best Local Similarity 61.7%; Pred. No. 4.4e-67;
Matches 490; Conservative 0; Mismatches 298; Indels 6; Gaps 3;

QY 1441 TTTTTTTTTTTTTTTTTTGAAGAGTCTGTCTTTTCACTCCCTGGCTGGAGGCGAGTGG 1500
Db 126376 TCTCTCTCTTTTCTTTTCAAGACAGAGTCTTGTCTCTCTCACCAGGCTGGAGTCA 126434

QY 1501 TGATCTCAGCTCAGTCACTTTCGCTCCCGGGTTCAAGCAATTCCTCTGCTTCAGCCT 1560
Db 126435 CGACCTTGCTCAGTCACTTCCGCTCCCTGCTTCTGGGTTCAAGAGATTCCTCTGCTCAGCTC 126494

QY 1561 CCCAAGTAGCTGGGACTCAGGCACTTCCACCATGCCCGAGCTGATTTTGTATTATTAG 1620
Db 126495 CCCGAGTAGCTGGGATTACAGTGCCCGCCACCATGCTGCTAATTTTGTATTATTAG 126554

QY 1621 TAGAGATGGGATTTCACTTTGTGTGCCAAGCTGGTCTCAAACTTTTGTCTGTCAATTG 1680
Db 126555 TAGAGATGGGTTTCCACCATGTGAAGCATTTTAGGAAAGGGGTTAGGAGGAGGAGACA 126614

QY 1681 TTGTAACTATTGTTCTTTTGTGAGGTAGGCGCCCGCCAGACCAAAAAATAAATCTTAG 1740
Db 126615 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 126674

QY 1741 AATCCAAATCAGTGTGTGTGGTTTGACCACTGTCTACTTGAGAACCCACAGTGTGACCGGC 1800
Db 126675 NNTCTCCCCCTTCTCCCAAAATATGGCAGACTCTTCTGTCTCCCTAGTCTCATATA 126734
```

QY 1801 CTCAGAGTAGAGTGATCTCTGCTCGAAGAGAAATAGAAATTAATCTTCGGGGCC 1860
 Db 126735 GTCCTCTGAATGTTTATTTCTTCCCTGTTAAATACAGTAAATAATTAAGTTGGCC 126794
 QY 1861 AGGCGTGGTGGCTCATGCTGTAATCCAGCATTTCAGGAGGCAAGGATGTGATCAC 1920
 Db 126795 AGGAGTGGTGGCACAATGCTGTAATCCAGCATTTCAGGAGGCGGAGGAGATCAC 126854
 QY 1921 CTGAGTCAAGGAGTTCACAAACAGCCTGGCCAAACATGGTGAACCCCGTCTCTACTAAAA 1980
 Db 126855 CTGAGTCAAGGAGTTCAGACCAAGCCTGGCCAAACATGGTGAACCCCGTCTCTACTAAAA 126914
 QY 1981 ATACA----AAAAATPAGCTAAGTGTGGTGGCGCATGCTCTGTAATCCAGCTACTTTGGGA 2036
 Db 126915 AAAAATGTAAAGTTAGCTGGGTGTGGTGGCGCATGCTCTGTAATCCAGCTACTTTGGGA 126974
 QY 2037 GGGTGAAGGAGGAGTTCCTCAACCCGGGAGGAGAGTTCAGTGAAGCGAGATCAC 2096
 Db 126975 GGGTGAAGGAGGAGTTCCTCAACCCGGGAGGAGAGTTCAGTGAAGCGAGATCAC 127034
 QY 2097 ACCACTGCACTCCAGCCTGGGGAGAGAGAGAGTTCCTCTCAAAAAACAAAAACAA 2156
 Db 127035 GCCACTGCACTTAGCTGGGAGAGAGAGTTCCTCTCAAAAAACAAAAACAA 127093
 QY 2157 AAGAATTAAGCAAAATPAGATTCAGAGAGAACCTGAAGGGGGTGCAGACGATACAGA 2216
 Db 127094 AGTTAGTAGTACTTTGGGCGCTATCAGATAGTGTCTCTCTGCGAGGCGGCTTGGGCGCCCTA 127153
 QY 2217 TTTCTGTGCCACAT 2230
 Db 127154 CTACAGTTTCACTT 127167

RESULT 9
 US-10-238-709-3
 ; Sequence 3, Application US/10238709
 ; Patent No. 6680188
 ; GENERAL INFORMATION:
 ; APPLICANT: WEBSTER, Marlon et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001164DIV
 ; CURRENT APPLICATION NUMBER: US/10/238,709
 ; CURRENT FILING DATE: 2002-09-11
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 174493
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(174493)
 ; OTHER INFORMATION: n = A,T,C or G

Query Match 7.9%; Score 312.4; DB 4; Length 174493;
 Best Local Similarity 61.7%; Pred. No. 4.4e-67;
 Matches 490; Conservative 0; Mismatches 298; Indels 6; Gaps 3;

QY 1441 TTTTCTTTTCTTTTGAAGAGATCTTCTTTCACCTGGCTGGAGGAGTGGTG 1500
 Db 126376 TCTCTTCTTTTCTTTCAAGACAGATCTTCTCTGTCTCACCCAGCTGGAGTGGAGTG-G 126434
 QY 1501 TGATCTCAGCTCACTGCAATTTTGGCTCCCGGGTCAAGCAATTCCTCTGCTTCAGCCT 1560
 Db 126435 CGACCTTGGCTCACTGCAACCTCCGCTCTCTGGTTCAGAGATTCCTCTGCTTCAGCTC 126494
 QY 1561 CCCAAGTAGCTGGAGCTACAGGAGCTTCCACCATGCCAGCTGATTTTGTATTTTAG 1620
 Db 126495 CCCGAGTAGCTGGAGTACAGGTGGCGCCGCCACCATGCTGGCTAATTTTGTATTTTAG 126554

QY 1621 TAGAGATGGGATTTCACTTTGTGTGGCCAAAGCTGGTCTCAAACTTTTGTGTCTAATTTG 1680
 Db 126555 TAGAGATGGGTTTCAACCATTTGAAGCATTTTAGAAAGGGGTTAGGAGGAGGAAGACA 126614
 QY 1681 TTGTAACTATTCTTCTTTTCTGCTGAGGTAGGGCCCCCAGACCAAAAAATAAATCTTAG 1740
 Db 126615 NNN 126674
 QY 1741 AATCCAAATCACTGTGTGTGTGTGACCACTGTCTCTGAGAACCAACAGTGTGACCAAGGC 1800
 Db 126675 NNTCTCCCCCTTTCCCAAAATATGGCAGACTTCTCTCTCCCTAGTCTTCTCATATA 126734
 QY 1801 CTCAGAGTAGAGGTGATCTCTGCTCGAAGAGAAATAGAAATAAATATTTCTCCGGGCC 1860
 Db 126735 GTCTTCTGAATGTTGTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 126794
 QY 1861 AGGCGTGGTGGCTCATGCTGTAATCCAGCATTTCAGGAGGCGCAAGGATGTGGATCAC 1920
 Db 126795 AGGAGTGGTGGCACAATGCTGTAATCCAGCATTTCAGGAGGCGGAGGAGATCAC 126854
 QY 1921 CTGAGTCAAGGAGTTCACAAACAGCCTGGCCAAACATGGTGAACCCCGTCTCTACTAAA 1980
 Db 126855 CTGAGTCAAGGAGTTCAGACCAAGCCTGGCCAAACATGGTGAACCCCGTCTCTACTAAA 126914
 QY 1981 ATACA----AAAAATPAGCTAAGTGTGGTGGCGCATGCTCTGTAATCCAGCTACTTTGGGA 2036
 Db 126915 AAAAATGTAAAGTTAGCTGGGTGTGGTGGCGCATGCTCTGTAATCCAGCTACTTTGGGA 126974
 QY 2037 GGGTGAAGGAGGAGTTCCTCAACCCGGGAGGAGAGTTCAGTGAAGCGAGATCAC 2096
 Db 126975 GGGTGAAGGAGGAGTTCCTCAACCCGGGAGGAGAGTTCAGTGAAGCGAGATCAC 127034
 QY 2097 ACCACTGCACTCCAGCCTGGGGAGAGAGAGAGTTCCTCTCAAAAAACAAAAACAA 2156
 Db 127035 GCCACTGCACTTAGCTGGGAGAGAGTTCCTCTCAAAAAACAAAAACAA 127093
 QY 2157 AAGAATTAAGCAAAATPAGATTCAGAGAGAACCTGAAGGGGGTGCAGACGATACAGA 2216
 Db 127094 AGTTAGTAGTACTTTGGGCGCTATCAGATAGTGTCTCTCTGCGAGGCGGCTTGGGCGCCCTA 127153
 QY 2217 TTTCTGTGCCACAT 2230
 Db 127154 CTACAGTTTCACTT 127167

RESULT 10
 US-09-345-882-1/c
 ; Sequence 1, Application US/09345882
 ; Patent No. 6399373
 ; GENERAL INFORMATION:
 ; APPLICANT: Bougueleret, Lydie
 ; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
 ; TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
 ; FILE REFERENCE: GENSET.031A
 ; CURRENT APPLICATION NUMBER: US/09/345,882
 ; CURRENT FILING DATE: 1999-06-30
 ; PRIOR APPLICATION NUMBER: US 60/091,315
 ; PRIOR FILING DATE: 1998-06-30
 ; PRIOR APPLICATION NUMBER: US 60/111,909
 ; PRIOR FILING DATE: 1998-12-10
 ; NUMBER OF SEQ ID NOS: 140
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 1
 ; LENGTH: 162450
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: allele
 ; LOCATION: 72794
 ; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
 ; FEATURE:
 ; NAME/KEY: allele

```
/ LOCATION: 88073
/ OTHER INFORMATION: 5-127-261 : polymorphic base A or C
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 90842
/ OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 93714
/ OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 97122
/ OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 97152
/ OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 99098
/ OTHER INFORMATION: 5-130-257 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 99117
/ OTHER INFORMATION: 5-130-276 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 103806
/ OTHER INFORMATION: 5-131-395 : polymorphic base A or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 106940
/ OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 108106
/ OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 108149
/ OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 108308
/ OTHER INFORMATION: 5-135-357 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 108471
/ OTHER INFORMATION: 5-136-174 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 134134
/ OTHER INFORMATION: 5-140-120 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 134362
/ OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 134374
/ OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 146328
/ OTHER INFORMATION: 5-143-84 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 146345
/ OTHER INFORMATION: 5-143-101 : polymorphic base A or C
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 150329
```

```
/ OTHER INFORMATION: 5-145-24 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 160031
/ OTHER INFORMATION: 5-148-352 : polymorphic base G or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 72771..72817
/ OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 72771..72817
/ OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 88050..88096
/ OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 88050..88096
/ OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 90819..90865
/ OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 90819..90865
/ OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 93690..93736
/ OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 93690..93736
/ OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 97099..97145
/ OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 97099..97145
/ OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 97130..97177
/ OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 97130..97177
/ OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 99075..99121
/ OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 99075..99121
/ OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 99094..99140
/ OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 99094..99140
/ OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 103783..103828
/ OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
```

1442	QY	TTTTTTTTTTTTTTTTGAGAAAGAGCTTGTGCTCTTTCACTCCCTGGCTGGAGGGCAGGGTGT	1501
9962	Db	TTTTTTTGTGTGTGAGACAGAGCTTGTCTCCGTCACTCCAGGATGGAGTGCAGTGGCGT	10021
1502	QY	GATCTCAGCTCACTGCAACTTTTCGGCTCCCGGTTCAAGCAATTCCTCGTCTCAGCCCTC	1561
10022	Db	GGCTTTGGCTCACTCGAACCTCTACTCTCTGGTTCAAGCAATTCCTCGTCTCAGCCCTC	10081
1562	QY	CCAAGTAGCTGGACTACAGGCACCTCCACCATGCCAGCTGATTTTGTATTTTATGT	1621
10082	Db	CCAAATAGCTGGGATTACAGGTGCGCACCAACACTCTCTGGCTAATTTTGTATTTTATGT	10141
1622	QY	AGAGATGGGATTTCACTTTGTGTGGCCAGCTGGCTCAAACTTTTGTGCTGCATAAATGT	1681
10142	Db	--AGACGGGGTTTTCACCATGTGTGGCCAGGCTGGTTTCAAGTAATCCACCCCTCAGCGCT	10199
1682	QY	TGTAACATATTGTTCCCTTTGTCTAGGTAGGGCCCCCAGACCAAAAAATAAATCTTAGA	1741
10200	Db	CCCCAAGTCTGGGATTACAGACTAGGCCACTGGCTGGSCCCCATCGGCCCTCCTATTT	10259
1742	QY	ATCCAAATCAGTGTGTGTGTGACCATGTGCACTGTGAGAACCAACAGTGTGACCAAGGCC	1801
10260	Db	TGTATGTGTGTGCCCTGCAATTAGAGCCCATATCTTGGATGTTTCCATTTGGGTATTAGGTCT	10319
1802	QY	TCAGAGTAGAGGTGATCTGTCTCGAAGAGAAATAGAAATGAAATATTCTCCGGGCCA	1861
10320	Db	GAGACAGCATCTTAGTCCGTGGGTGCCACGCTTGTCAGAAATCTTGATTTCTGGGCCA	10379
1862	QY	GGCGTGTGGGTCACTGCTGTAAATCCACGACTTTGGGAGGCCAAGGCATGTGGATCACC	1921
10380	Db	GGCATGGTGGGTCACTCTGTAAATCCACGACTTTGGGAGGCCAAGGCATGTGGATCA--	10437

QY 1922 TGAGGTGAGGATTCAAAACCCAGCTGGCCAAACATGGTGAACCCCGTCTCTACTAAAAA 1981
Db 10438 TGAGGTGAGGATTAGAGACACAGCTGGCCAAACATGGTGAACCCCTGTCTACTAAAAA 10497
QY 1982 TACAAAAAATTAGCTAAGTGTGGTGGCGCATGCTCTAATCCAGCTACTTGGGAGGGTG 2041
Db 10498 TAGAAA-AAATTAGCTGGGTGGTGGCGGTACTATAATCCAGCTACTTGGGAGGGCTG 10556
QY 2042 AGGAGGAGAAATTTCTTGAACCCG-GGAGGAGAGGTTGAGTGAAGCGAGATCACACCA 2100
Db 10557 AGGAGGAGATCATTTGAACCTGAGGGGTGAGGTTGAGTGAAGCGAGATCATACCA 10616
QY 2101 CTGCACTCCAGCTGGGAGAGAGAGAGACTTCTCTCAAAAAAACAACAAAAAAGA 2160
Db 10617 TTGCACTCCAGCTGGGTGACAGGGTGAGACTCCGTCTCAAAAAAACAACAAAAAAGA 10676
QY 2161 ATTAAG 2166
Db 10677 ATCCAG 10682
RESULT 12
US-10-060-332-3
; Sequence 3, Application US/10060332
; Patent No. 6528294
; GENERAL INFORMATION:
; APPLICANT: LI, JIAYIN et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE OF INVENTION: THEROP
; FILE REFERENCE: CL000851DIV
; CURRENT APPLICATION NUMBER: US/10/060,332
; CURRENT FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 43950
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-332-3

Query Match 7.8%; Score 307.6; DB 4; Length 43950;
Best Local Similarity 67.6%; Pred. No. 3.4e-66;
Matches 491; Conservative 0; Mismatches 229; Indels 6; Gaps 4;
QY 1442 TTTTTTTTTTTTTTTGAGAAAGAGTCTTGCTCTTTACACCTGGCTGGAGGCGAGTGGT 1501
Db 9962 TTTTTTTGTGTGTGAGACAGATCTTGCTCGTCAACCCAGATGGAGTGGAGTGGGT 10021
QY 1502 GATCTAGCTCACTGCAACTTTGGCTCCCGGTTCAAGCAATTCCTGCTTCAGCCTC 1561
Db 10022 GGTCTGGCTCACTGCAACTTACCTCTGCTGGTTCAAGCAATTCCTGCTTCAGCCTC 10081
QY 1562 CCAAGTAGCTGGGATACAGGCACTTCCACCATGCCAGCTGATTTTGTATTTTGTAGT 1621
Db 10082 CCAATAGCTGGGATACAGGCGGATCAGGCGGATCAGGCGGATTTTGTATTTTGTAGT 10141
QY 1622 AGAGATGGGATTTCACTTTGTTGGCAAGCTGGTCTCAAACTTTTGTGTCATAATTGT 1681
Db 10142 --AGACGGGTTTACCATGTTGGCCAGGCTGTTTCAAGTAATCCACCTCCTCAGCT 10199
QY 1682 TGTAATATTGTTCTTTTGTGAGTAGGGGCCCCAGACCAAAAAAATAATCTTAGA 1741
Db 10200 CCCCAGGTGCTGGGATTAAGACATGAGCCAATGGGCTGGGCCCATCGGCTCCATTTT 10259
QY 1742 ATCCAAATCAGTGTGTGGTTTGAACCTGTCACTTGAGAACCAAGTGTGACAGGGCC 1801
Db 10260 TGATGTTGTCCTCCGCAATTAAGACCAATTTCTTGGATGTTCCATTTGGTATTAGGTT 10319
QY 1802 TCAGGAGTAGAGTGATCTCTCTCGAAAGAGAAATAGAAATGAAATATTTCTCCGGGCCA 1861
Db 10320 GAGACAGCATCTAGTCCGTGGGTGCCACGCTTGTACAGAAATCTTGATTTCTGGGCCA 10379

QY 1862 GGCGTGGTGGCTCATGCTGTATATCCAGCACTTTGGGAGGCCAAGGCAATGTGATCAAC 1921
Db 10380 GGCAAGTGGCTCACACCTGTATATCCAGCACTTTGGGAGGCCAAGGCGGCGATCA-- 10437
QY 1922 TGAGGTGAGGATTCAAAACCCAGCTGGCCAAACATGGTGAACCCCGTCTCTACTAAAAA 1981
Db 10438 TGAGGTGAGGATTAGAGACACAGCTGGCCAAACATGGTGAACCCCTGTCTACTAAAAA 10497
QY 1982 TACAAAAAATTAGCTAAGTGTGGTGGCGCATGCTCTAATCCAGCTACTTGGGAGGGTG 2041
Db 10498 TAGAAA-AAATTAGCTGGGTGGTGGCGGTACTATAATCCAGCTACTTGGGAGGGCTG 10556
QY 2042 AGGAGGAGAAATTTCTTGAACCCG-GGAGGAGAGGTTGAGTGAAGCGAGATCACACCA 2100
Db 10557 AGGAGGAGATCATTTGAACCTGAGGGGTGAGGTTGAGTGAAGCGAGATCATACCA 10616
QY 2101 CTGCACTCCAGCTGGGAGAGAGAGAGACTTCTCTCAAAAAAACAACAAAAAAGA 2160
Db 10617 TTGCACTCCAGCTGGGTGACAGGGTGAGACTCCGTCTCAAAAAAACAACAAAAAAGA 10676
QY 2161 ATTAAG 2166
Db 10677 ATCCAG 10682
RESULT 13
US-09-851-896-3/c
; Sequence 3, Application US/09851896
; Patent No. 6410325
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPEND
; FILE REFERENCE: RTS-0220
; CURRENT APPLICATION NUMBER: US/09/851.896
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 70000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-851-896-3

Query Match 7.8%; Score 307.2; DB 4; Length 70000;
Best Local Similarity 70.2%; Pred. No. 5.4e-66;
Matches 504; Conservative 0; Mismatches 193; Indels 21; Gaps 6;
QY 1441 TTTTTTTTTTTTTTTGAGAAAGAGTCTTGCTCTTTTCAACCTGGCTGGAGGCGAGTGGTG 1500
Db 50400 TTGTTTTTTTTTTTTTTGAGACGGAGTCTCACTCTGTCAACCCAGGCTGGAGTGGCG 50341
QY 1501 TGATCTAGCTCACTGCAACTTTCGCTCCCGGTTCAAGCAATTCCTGCTTCAGCCT 1560
Db 50340 TGATCTAGCTCACTGCAACCTCTACCTACCGAGTTCAAGCAATTCCTGCTTCAGCCT 50281
QY 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCACCATGCCAGCTGA-TTTTGTATTTTGA 1619
Db 50280 CCAGATAGCTGGTATTACAGTCCCAACCACTACGCCAGCTAAATTTTGTATTTTGA 50221
QY 1620 GTAGATGGGATTTCACTTTGTTGGCAAGCTGGTCTCAAACTTTTGTCTGTCTATATT 1679
Db 50220 GTAGAGACAGGTTTCAACATGTTGGCCAGGCTGTCTTGAC---TCCTGACCTCGTG 50165
QY 1680 GTTGTAACTATTGTTCTTTGCTGAGTAGGGCCCCAGACCAAAAAAATAAATCTTA 1739
Db 50164 ATTCCGCCACCTCCGGCTCCCAAAATGCTGGATTACAGGCGTGAGCCACCGCGCCT-- 50107
QY 1740 GAATCCAAATCAGTGTGTGTTTGACCACTGTCTACATTGAGAACCAAGTGTGACAGGG 1799

```

Db 50106 -----GCAATGGTTAAGCTTTTATGGTTGGCACT--GGATGGATGGATGGCCATTTA 50058
QY 1800 CCTCAGGAGTAGAGGTGATCTCTGCTCGAAGAGAAATAGAAATGAAATATTTCTCCGGGC 1859
Db 50057 TGGTAGTGATGCAACATTTCTCTCAAAATTAATTAAGTCAATTTAAAGTAGCCAGGC 49998
QY 1860 CAGCGTGGTGGCTCATGCTCTGTAATCCAGCACTTTGGGAGGCCAAGGCATGGATCA 1919
Db 49997 CGGCAAGGTGGCTCAGCTTGTATATCCAGCACTTTGGGAGGCTGAGGCAGGTGGAACA 49938
QY 1920 CCTGAGTCAAGGATTCAAAACAGCCTGGCCACATGGTGAACCCCGTCTCTACTAAA 1979
Db 49937 CCTGAGTCAAGGATTCAAAACAGCCTGGCCACATGGTGAACCCCGTCTCTACTAAA 49878
QY 1980 AATACAAAAAATAGCTTAAGTGGTGGCGCATGCCCTGTAATCCAGCTACTTTGGGAGGG 2039
Db 49877 AATAC--AAATTAAGCAGGTGGTGGCGCATGAACGTAATCCAGCTACTCTGGGAGC 49820
QY 2040 TGAGGCAGGAGAAATTTCTCAACCCGGGAGCGAGAGTTGCAAGTGAAGCAGATCAACCC 2099
Db 49819 TGAGGCAGGAGAAATCGCTTGAACCCAGGAGCGAGGTTGGGTTGAGCCAAAGATCGTGCC 49760
QY 2100 ACTGCATCCAGCCTGGGGA--GAGAGCGAGACTTCTCTCAAAAACAAAAACAA 2156
Db 49759 ATTGCATCCAGCCTGGGCAACAAAAGCGAAACTCTCTCAAAATAAATAAATAAAA 49702

RESULT 14
US-09-820-924-3/c
; Sequence 3, Application US/09820924
; Patent No. 655351
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CU001213
; CURRENT APPLICATION NUMBER: US/09/820,924
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 39982
; TYPE: DNA
; ORGANISM: Human
US-09-820-924-3

Query Match 7.7%; Score 304; DB 4; Length 39982;
Best Local Similarity 56.3%; Pred. No. 2.5e-65;
Matches 829; Conservative 0; Mismatches 595; Indels 48; Gaps 12;

QY 735 TATTTTCTTTCTTGAACAGGCTATCACTCTGTGTCGCGAGCTGGAGTACAGTGGCGTAA 794
Db 32734 TATTAATTTTGTGATGAGTCTCATTTTGTGCGCCATGCTGGAGTGCAGTGCATGA 32675
QY 795 TAATAGCTCATCGACCTCC--CCTCTGGGCTCAAGCAATCCGCTGGCTCAGCATCCT 853
Db 32674 TCTTGGCTCACTACAACTCCACTCCCGGTTCAAGCTATTCTTGTGCTCAGCCTCC 32615
QY 854 GAGTACTGGGACTACAGGCTTGTGCAACAGCCAGCTAGTTTAAATAATGATTTT 913
Db 32614 GAGTACTGGGATTAAGGCAACCCACCATACCTACCTGGCTAATTTTTT-----GTACTTTT 32561
QY 914 TGTATAGAGGAGTCTTGTATGTGCTCAGGCTGATTTTTTATTTTGTGAGACAAGGTC 973
Db 32560 TAGTAGAGTGGGTTCCACATGTTGGCGGCTGGCTTGAATCTGTCTCAAGTGA 32501
QY 974 TCACTATGTTGCCATGATCCCCCACTCCACTTCCCAAGTGTCTCATC--TTATCTGTT 1031
Db 32500 TCCGCTGCTTTGGCTCCCAAGTGTGAGTATTAAGTGTGTGCCACCAAGCCAGCT 32441
QY 1032 CATTAGTCAAGTCAACAGATTTAGGTTGTTTCCACTTTTGGACCAATATGATTAATCT 1091

```

```

Db 32440 AATTTTGTATTTTATAGTGAGAAGGGGTTTACCATGTTGGCCAGGCTGCTCTCGAACT 32381
QY 1092 CCAGTGAATATTCATGTATACATTTGTGTGGGATATGTTTTCATTT----- 1138
Db 32380 CCTGACCTCTGTGATTCGGCGCTCAGCTCCCAAGTGTGGGATTAACAAGCCAGCC 32321
QY 1139 -----CTGTTGGGTTTATATCTAGGAGTGAATTCGTGGATCCCGGTAATATTTTGACA 1193
Db 32320 TGATACAGTGTGTTTAAATATGAGAGATCTCTACAAACCAATGTGGAATCATTTTCTTA 32261
QY 1194 GGCAGAGTTCCGGGAGAGAAAACTTTGGGAAAAATGAAGCATGTTTAG-----AAA 1243
Db 32260 ATAAATTTGTTAACTGAAGAAATGCAAAAGGAAGACAGAGTAGTCTAGGGTACTACTTCTT 32201
QY 1244 TCAGCAGAGTGCAGGGGTTTTTCGGAGTTTTTATTTTATATTCTGTTGACAATATGTCAG 1303
Db 32200 GCTGTGAGAAAGAGGGAAGATAGCATACTTGTGCTGCGCTGTTGTGCTAAACGAAC 32141
QY 1304 TTTGATGAAGATACAAAGTTATATCTAAGTGAAGATGAAGTAAAGGCTGGAATAGGGCGT 1363
Db 32140 CAGGAGGAGGAACCG--AAGTGAACGAGACTGCTTACCTACAAGGGGTAGGGGAGA 32083
QY 1364 TCAGAGTAAATCATGAGCACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAA 1423
Db 32082 AAACAGGGCAAAAGGAGAGTACTGTTATGTTAGTCTTCTGCTTTTGGAGGTACATAA 32023
QY 1424 ATAAATAAAATCACAATTTTTTTTTTTTTTTTTTTTGTGAGAAAGACTTGTCTTTTACCCTG 1483
Db 32022 TGTCTTTTATTTTTTATTTATTTATTTTATTTTATTTTATGACAGAGTCTCACTGT--TGCAG 31964
QY 1484 GCTGAGGCGAGTGTGTGATCTCAGCTCACTCAACTTTTCGCTCC--GGGTTCGAAGCA 1542
Db 31963 GCTGGAGTGCAGTGTGTGATCACACTCACTGAGCTCGACTCCAGGGCTCAGGTG 31904
QY 1543 ATTCTCTGCTCAGCTCCCAAGTGTGGGACTCAGGCACTTCCACCATGCCCCAGC 1602
Db 31903 ATCTCCCACTCAGCTCCAGGTTAGTGGGACTTCAAGTGTGCTACCAACCCAGC 31844
QY 1603 TGATTTTCTATTTTATAGTAGAGTGGGATTTTCACTTTTGGCCAGCTGGTCTCAAC 1662
Db 31843 TAATTTTGTATTTCTTTGTAGACA--GAGGTTTGGCCATGTTGCCAGGCTGGTCTCAAC 31785
QY 1663 T-TTTTGTCTCATAAATTTGTGTAACTATTTGTTCTTTTGTGAGTGGGCCCCAGAC 1721
Db 31784 TCCTGGGCTCAAGCAATCCATCTGCTTGGCTCCCAAGTGTAGGATTTACAGGCATGA 31725
QY 1722 CAAAAAATAATCTTAGAATCCAAATCAGTGTGTGGTTTGACCACTGTCTACTTGAGA 1781
Db 31724 GACACTGTGCCAGCCACCGTTGTTTCAAAACATTCAAAACAACAGCAACAATAACAACG 31665
QY 1782 ACCACAGTGTGACCCAGGGCTCAGGAGTAGAGTGTATCTCTGCTCGAAAGAGAAATAGAA 1841
Db 31664 CTCAAACCAAAACAGAAATCAGCAAGATGAGAGGGAATCCCTAAGTGGATTAATAA 31605
QY 1842 TGAATAATTTCTCCGGCCAGCGTGGTGGCTCAGCTGTAAATCCAGCACTTTGGGAG 1901
Db 31604 TGAATCCAA---CTAGGCTGGGCGGCTGCTCAGCTGTAAATCCAGCACTTTGGGAG 31548
QY 1902 GCCAGGCAATGGAATCAGTGTGAGTCAAGGTTCAAAACCCAGCTGGCCAAACATGGTGA 1961
Db 31547 GCGAGGCGGGCGGATTCAC--GAGGTCAGGAGATCGAGCCATCTCTGGGTAAACCATGA 31490
QY 1962 AACCCCGCTCTACTAAAAAATACAAAAATTAGCTAAGTGTGGTGGCGCATGCTCTAAT 2021
Db 31489 AACCCCATCTCTACTAAAAAATACAAAAATTAGCCGGGCTGGTAGCGGCACCTAGT 31430
QY 2022 CCAGCTACTGGGAGGTGAGCGAGGAAATTTCTTGAACCCGGAGGAGCAGAGTTGCA 2081
Db 31429 CCAGCTACTGGGAGGTGAGCGAGGAAATGGCGTGAACCCGGAGGCGAGAGCTTGA 31370
QY 2082 GTGAAGCGAGATCACACACTGCTCCAGCTGGGGGAGAGCGAGACTTCTCTCAA 2141
Db 31369 GTGAGCTGAGTAGTGCCACTGCACTCCGCGCTGGGCGAAAGAGCAAGACTCCGCTCAA 31310

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 09:06:53 ; Search time 2270.77 Seconds
(without alignments)
7909.918 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3958

Perfect score: 3958

Sequence: 1 aggaattcatcatttaaat.....gggacctcagcaccgcgg 3958

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 2947324 seqs, 2269024515 residues

Total number of hits satisfying chosen parameters: 5894648

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3958	100.0	4150	14	US-10-016-725-15
2	3929.8	99.3	37698	13	Sequence 15, Appl
3	630.6	15.9	631	13	Sequence 10, Appl
4	630.6	15.9	631	16	US-10-021-632-24494
5	498.4	12.6	617	10	Sequence 24494, A
6	460	11.6	59725	13	Sequence 1, Appl
7	407.4	10.3	9631	15	US-09-792-4698-1
8	360.4	9.1	23618	13	Sequence 814, App
9	353.6	8.9	36296	13	Sequence 1569, Ap
10	353	8.9	60815	13	Sequence 4, Appl
11	349.8	8.8	12394	16	US-10-240-425-1584
12	349	8.8	13216	10	US-10-087-192-52
13	347.4	8.8	29163	10	US-10-181-875-10
14	342.8	8.7	26928	9	US-09-764-891-7810
					Sequence 7809, Ap
					Sequence 2278, Ap

15	342.8	8.7	26928	13	US-10-235-192A-34	Sequence 34, Appl
16	342.8	8.7	26928	15	US-10-020-141-7	Sequence 7, Appl
17	342.8	8.7	26928	15	US-10-017-631-1	Sequence 1, Appl
18	338.6	8.6	144035	13	US-10-087-192-322	Sequence 322, App
19	337.2	8.5	3273	16	US-10-012-697-1406	GENERAL INFORMATI
20	337	8.5	109906	13	US-10-235-192A-31	Sequence 31, Appl
21	334.8	8.5	108182	13	US-10-087-192-1618	Sequence 1618, Ap
22	333.8	8.4	175077	13	US-10-087-192-1168	Sequence 1168, Ap
23	333.6	8.4	35425	15	US-10-017-161-2429	Sequence 2429, Ap
24	333.6	8.4	35425	16	US-10-292-798-2069	Sequence 2069, Ap
25	332.2	8.4	3158	13	US-10-027-632-115030	Sequence 115030, Ap
26	332.2	8.4	3158	16	US-10-027-632-115030	Sequence 115030, Ap
27	330.6	8.4	3158	13	US-10-027-632-115030	Sequence 115030, Ap
28	330.6	8.4	3158	16	US-10-027-632-115030	Sequence 115030, Ap
29	330.2	8.3	198522	13	US-10-087-192-244	Sequence 244, App
30	329.6	8.3	9372	9	US-09-764-877-3049	Sequence 3049, Ap
31	329.6	8.3	9372	9	US-09-764-877-3244	Sequence 3244, Ap
32	329.6	8.3	9372	16	US-10-242-515-3049	Sequence 3049, Ap
33	329.6	8.3	9372	16	US-10-242-515-3244	Sequence 3244, Ap
34	328.8	8.3	2130	16	US-10-108-260A-2246	Sequence 2246, Ap
35	328.8	8.3	75252	13	US-10-087-192-904	Sequence 904, App
36	328.6	8.3	35425	15	US-10-017-161-2429	Sequence 2429, Ap
37	328.6	8.3	35425	16	US-10-292-798-2069	Sequence 2069, Ap
38	328.4	8.3	96594	12	US-09-997-722-154	Sequence 154, App
39	328	8.3	91760	13	US-10-087-192-844	Sequence 844, App
40	327.6	8.3	12919	15	US-10-017-161-787	Sequence 787, App
41	327.6	8.3	17397	9	US-09-764-869-1945	Sequence 1945, Ap
42	327.6	8.3	17397	15	US-10-091-504-1945	Sequence 1945, Ap
43	327.6	8.3	17397	16	US-10-227-577-1945	Sequence 1945, Ap
44	327.6	8.3	19334	9	US-09-764-869-1943	Sequence 1943, Ap
45	327.6	8.3	19334	15	US-10-091-504-1943	Sequence 1943, Ap

ALIGNMENTS

RESULT 1

US-10-016-725-15
; Sequence 15, Application US/10016725
; Publication No. US20020151018A1
; GENERAL INFORMATION:
; APPLICANT: Prouty, Stephen
; APPLICANT: Zhang, Lin
; APPLICANT: Stenn, Kurt
; TITLE OF INVENTION: Stearyl-CoA Desaturase Promoter
; FILE REFERENCE: J&J2065
; CURRENT APPLICATION NUMBER: US/10/016,725
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 15
; LENGTH: 4150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-016-725-15

Query Match 100.0%; Score 3958; DB 14; Length 4150;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AGGAATTCATCCATTTAAATCATACATTAATTAATGCGTTTATAGTATATTCACAGGTTGTC	60
Db	1	AGGAATTCATCCATTTAAATCATACATTAATTAATGCGTTTATAGTATATTCACAGGTTGTC	60
QY	61	ATCCATCAATCCATTTTAGAACAGTTTATTAATCTCCAAAAATAAACCTGCATTCCTT	120
Db	61	ATCCATCAATCCATTTTAGAACAGTTTATTAATCTCCAAAAATAAACCTGCATTCCTT	120
QY	121	AGCATACACCCCAACATCCCTCCATTCCTTCAGCCCTGGGCAACCAACCAATCTAC	180
Db	121	AGCATACACCCCAACATCCCTCCATTCCTTCAGCCCTGGGCAACCAACCAATCTAC	180
QY	181	TTTCTGCTCTATTAATTTGCAATTTCTGGACATTTTCATATAAATGGAAGCAACCAACAT	240

Db 181 TTTCTGCTCTATATAAATTTGCCAATTTCTGGCAATTTTCATATAAAATGGAAGCAAAACAACAT 240
Qy 241 GTGAGACTTTGTGACTGCTGCTGTTTCACTTAGCATTTCTATTTTAAAGGCTCATATATGTTA 300
Db 241 GTGAGACTTTGTGACTGCTGCTGTTTCACTTAGCATTTCTATTTTAAAGGCTCATATATGTTA 300
Qy 301 CAGTACTTAGCAGTACTTCACTTTCTTTTATTTCTCAAATGATTTTCCACTGTGTGGGTAT 360
Db 301 CAGTACTTAGCAGTACTTCACTTTCTTTTATTTCTCAAATGATTTTCCACTGTGTGGGTAT 360
Qy 361 CCCATATCATATATTAGACAGAGTCTCTCACTGTGTCAACCCAGGCTGAGTGCAGTGGC 420
Db 361 CCCATATCATATATTAGACAGAGTCTCTCACTGTGTCAACCCAGGCTGAGTGCAGTGGC 420
Qy 421 ACAATCATAGCTCACTGTAACTCAAACTCCCTGGGCTCAAGTGATCTTACTTACCTCAGCC 480
Db 421 ACAATCATAGCTCACTGTAACTCAAACTCCCTGGGCTCAAGTGATCTTACTTACCTCAGCC 480
Qy 481 TCCAGAGTAGCTAGGACTACAGGCAACACAGCCATACTCGCTAAATTTTTTTTTTAAT 540
Db 481 TCCAGAGTAGCTAGGACTACAGGCAACACAGCCATACTCGCTAAATTTTTTTTTTAAT 540
Qy 541 TTTCAATTTTATGATATTCATTTTCTTTCTTTTGTGTTGTTGTTTGTAGATAGGCTCTC 600
Db 541 TTTCAATTTTATGATATTCATTTTCTTTTGTGTTGTTGTTTGTAGATAGGCTCTC 600
Qy 601 ACTTTGTTTACCCAGGCTGGAGGCACTGGCATGGTGACAGCTGAGCAGCTTGACTTCCT 660
Db 601 ACTTTGTTTACCCAGGCTGGAGGCACTGGCATGGTGACAGCTGAGCAGCTTGACTTCCT 660
Qy 661 GGGCTCAAGTGAATCTCTGCTCAGCCTTCCAGTGTCCAGTGTGGGACTA CAACAACGTGTCA 720
Db 661 GGGCTCAAGTGAATCTCTGCTCAGCCTTCCAGTGTCCAGTGTGGGACTA CAACAACGTGTCA 720
Qy 721 CCATGCTGCTGTATTTTCTTGAACAGGTTATCACTGTTGCCAGCTGGA 780
Db 721 CCATGCTGCTGTATTTTCTTGAACAGGTTATCACTGTTGCCAGCTGGA 780
Qy 781 GTACAGTGGCGTAAATAGCTCACTGACAGCTCCCTCCCTGGGCTCAAGCAATCCGCTG 840
Db 781 GTACAGTGGCGTAAATAGCTCACTGACAGCTCCCTCCCTGGGCTCAAGCAATCCGCTG 840
Qy 841 GCCTCAGCATCTCTAGTAGTGGGACTACAGCTTGTGCGCACAGGCTCAGCTAAGTATT 900
Db 841 GCCTCAGCATCTCTAGTAGTGGGACTACAGCTTGTGCGCACAGGCTCAGCTAAGTATT 900
Qy 901 AAAAAATGATTTTGGTATAGGAGGCTTGTCTATGTTGCTCAGGCTGATTTTATTTG 960
Db 901 AAAAAATGATTTTGGTATAGGAGGCTTGTCTATGTTGCTCAGGCTGATTTTATTTG 960
Qy 961 TTGAGCAAGGCTCTCACTATGTTGCCATGATCCCCCACCCTCCACTTCCCAGAGTCTCA 1020
Db 961 TTGAGCAAGGCTCTCACTATGTTGCCATGATCCCCCACCCTCCACTTCCCAGAGTCTCA 1020
Qy 1021 TCTTATCTGTTCAATTAGTCAGTTGACAGCAATTTAGTGTGTTTCACTTTTGAACCAATTA 1080
Db 1021 TCTTATCTGTTCAATTAGTCAGTTGACAGCAATTTAGTGTGTTTCACTTTTGAACCAATTA 1080
Qy 1081 TGAATAATCTCCAGTGAATATTCATGTATACATTTGTGGGCATATGTTTCAATTTCT 1140
Db 1081 TGAATAATCTCCAGTGAATATTCATGTATACATTTGTGGGCATATGTTTCAATTTCT 1140
Qy 1141 GTTGGGTTTATATCTAGGAGTGAATTTGCTGGATCCCGGTAATATTTTGAAGCAGAG 1200
Db 1141 GTTGGGTTTATATCTAGGAGTGAATTTGCTGGATCCCGGTAATATTTTGAAGCAGAG 1200
Qy 1201 TTCAGGGGAAGAAAACTTGGGAAAAATGAGCATGTTTAGAAATCAGCAAGTGCAGGG 1260
Db 1201 TTCAGGGGAAGAAAACTTGGGAAAAATGAGCATGTTTAGAAATCAGCAAGTGCAGGG 1260
Qy 1261 GTTTTTCGAGTTTATTTATTTCTGTTGACAAATGTGCAAGTTTGTATGAAGATACAAAG 1320
Db 1261 GTTTTTCGAGTTTATTTATTTCTGTTGACAAATGTGCAAGTTTGTATGAAGATACAAAG 1320

Db 1261 GTTTTTCGAGTTTATTTATTTCTGTTGACAAATGTGCAAGTTTGTATGAAGATACAAAG 1320
Qy 1321 TTATACTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380
Db 1321 TTATACTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380
Qy 1381 AGCACTTTGAATACCAAAATTAAGAGCTTGGCTGTAAACAAAAATAAATAAAAAATCAAA 1440
Db 1381 AGCACTTTGAATACCAAAATTAAGAGCTTGGCTGTAAACAAAAATAAATAAAAAATCAAA 1440
Qy 1441 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1500
Db 1441 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1500
Qy 1501 TGATCTCAGCTCACTGCAAACTTTCCGCTTCCCGGCTTCAAGCAATTTCTCTCTCTTCAGCCT 1560
Db 1501 TGATCTCAGCTCACTGCAAACTTTCCGCTTCCCGGCTTCAAGCAATTTCTCTCTCTTCAGCCT 1560
Qy 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCCACTCCCGGCTTCAAGCAATTTCTCTCTCTTCAGCCT 1620
Db 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCCACTCCCGGCTTCAAGCAATTTCTCTCTCTTCAGCCT 1620
Qy 1621 TAGAGATGGGATTTTCACTTTGTCGCAAGCTGTCTCAAACTTTTGTCTGTCAATAATTG 1680
Db 1621 TAGAGATGGGATTTTCACTTTGTCGCAAGCTGTCTCAAACTTTTGTCTGTCAATAATTG 1680
Qy 1681 TTGTAACATTTATTTTCTTTTCTGAGTGTAGGCTTCCAGACCAAAATAAATAAATAAATAAATAAATAA 1740
Db 1681 TTGTAACATTTATTTTCTTTTCTGAGTGTAGGCTTCCAGACCAAAATAAATAAATAAATAAATAAATAA 1740
Qy 1741 AATCAAAATCAGTGTGTTTGTGACCACTGTCTGAGAACCAACAGTGTGACAGGCT 1800
Db 1741 AATCAAAATCAGTGTGTTTGTGACCACTGTCTGAGAACCAACAGTGTGACAGGCT 1800
Qy 1801 CTGAGGAGTACAGTGTGATCTCTGCTCGAAAGAGAAATAGAAATGAAATATTTCTCCGGGCT 1860
Db 1801 CTGAGGAGTACAGTGTGATCTCTGCTCGAAAGAGAAATAGAAATGAAATATTTCTCCGGGCT 1860
Qy 1861 AGGCTGTGTGCTCATGCTTAATCCAGCACTTTGGGAGGCCAAGCATGTGGATCAC 1920
Db 1861 AGGCTGTGTGCTCATGCTTAATCCAGCACTTTGGGAGGCCAAGCATGTGGATCAC 1920
Qy 1921 CTGAGGCTCAGAGTTCAAAACAGGCTGGCCAAACATGTTGAAACCCCGTCTCTACTATAAA 1980
Db 1921 CTGAGGCTCAGAGTTCAAAACAGGCTGGCCAAACATGTTGAAACCCCGTCTCTACTATAAA 1980
Qy 1981 ATACAAAAATTTAGCTAAGTGTGTGCGCATGCTGTAATCCAGCTACTTGGAGGGT 2040
Db 1981 ATACAAAAATTTAGCTAAGTGTGTGCGCATGCTGTAATCCAGCTACTTGGAGGGT 2040
Qy 2041 GAGCAGGAGAAATTTCTTGAACCCGGGAGGAGGTTGCAAGTGAAGCAGATCACACCA 2100
Db 2041 GAGCAGGAGAAATTTCTTGAACCCGGGAGGAGGTTGCAAGTGAAGCAGATCACACCA 2100
Qy 2101 CTGCACTCCAGCTTGGGAGGAGAGGAGCTTCTCTCAAAAAAACHAAAAAACHAAAAA 2160
Db 2101 CTGCACTCCAGCTTGGGAGGAGAGGAGCTTCTCTCAAAAAAACHAAAAAACHAAAAA 2160
Qy 2161 ATTAAGCAAAATTTAGACATTTGACAGAGAACTGAAAGGGGGTCAACACAGTACAGATTTT 2220
Db 2161 ATTAAGCAAAATTTAGACATTTGACAGAGAACTGAAAGGGGGTCAACACAGTACAGATTTT 2220
Qy 2221 TGTGCCCATGCCAAGTACTTCTGAGGCACTGATGGAATGAGCTGTCCACATCTCTGAAATCA 2280
Db 2221 TGTGCCCATGCCAAGTACTTCTGAGGCACTGATGGAATGAGCTGTCCACATCTCTGAAATCA 2280
Qy 2281 TCCAGTCTTCTGTAGAACTTTTCAACCCGAGCAGGAGGCCAGGACTGGAATGAGTCTCT 2340
Db 2281 TCCAGTCTTCTGTAGAACTTTTCAACCCGAGCAGGAGGCCAGGACTGGAATGAGTCTCT 2340
Qy 2341 GGTCACTGGCCAGAGAGTGTGGCTTGAACCCCTGAGACAGTGGCCAAACAAAGAGAGCTCTT 2400
Db 2341 GGTCACTGGCCAGAGAGTGTGGCTTGAACCCCTGAGACAGTGGCCAAACAAAGAGAGCTCTT 2400

```
QY 2401 AGTCTACTCCCGAGAAATCCCGAGTCTGTGTCTTCTCTGGAAGTGAATCATTTGGGCGAG 2460
Db 2401 AGTCTACTCCCGAGAAATCCCGAGTCTGTGTCTTCTCTGGAAGTGAATCATTTGGGCGAG 2460
QY 2461 CACTCCGATATTTCTCTCTCTTCCAGGGGAGGATCCTAGGCGAGTATTTGGGAAGACA 2520
Db 2461 CACTCCGATATTTCTCTCTCTTCCAGGGGAGGATCCTAGGCGAGTATTTGGGAAGACA 2520
QY 2521 TGGGCATGGAAGGACACCGGGTGAATCATAGCTGCTGCTTCTGAGCTCTCATGTAA 2580
Db 2521 TGGGCATGGAAGGACACCGGGTGAATCATAGCTGCTGCTTCTGAGCTCTCATGTAA 2580
QY 2581 GGCTCTTACAGACACGAAAGATGAGGGGACAGGGAAGATGAGGCTGAGAGATC 2640
Db 2581 GGCTCTTACAGACACGAAAGATGAGGGGACAGGGAAGATGAGGCTGAGAGATC 2640
QY 2641 TCAGGACCGGCGCAATATGCTCTGAGCAGGATTAAGAGCTTGGGCTCTCATATGGT 2700
Db 2641 TCAGGACCGGCGCAATATGCTCTGAGCAGGATTAAGAGCTTGGGCTCTCATATGGT 2700
QY 2701 GTTCTGGGCTCAACTGCCAGCTCCGTCACTTACTGCTGTGACCATGGGCAAGTTA 2760
Db 2701 GTTCTGGGCTCAACTGCCAGCTCCGTCACTTACTGCTGTGACCATGGGCAAGTTA 2760
QY 2761 TTCATCTCTCCATATCTCTTTTCTCACTTTTAAATGGAATTAATGGGTACCCACCTCC 2820
Db 2761 TTCATCTCTCCATATCTCTTTTCTCACTTTTAAATGGAATTAATGGGTACCCACCTCC 2820
QY 2821 CAGGCTCACAGAGGCTTACAGAAACGATTTCTGTGAATTTGGCTTGGCATTAATTC 2880
Db 2821 CAGGCTCACAGAGGCTTACAGAAACGATTTCTGTGAATTTGGCTTGGCATTAATTC 2880
QY 2881 AATACCTGCCAGCTATTTCTTANTCCATCAAGCCCTTTCCGCTGCTGTGGTGAATA 2940
Db 2881 AATACCTGCCAGCTATTTCTTANTCCATCAAGCCCTTTCCGCTGCTGTGGTGAATA 2940
QY 2941 CATATGTCAGTGTCTCTGAGCTTTTCCACAAAGAGATTTCCAAATTAACAACCTGCCAG 3000
Db 2941 CATATGTCAGTGTCTCTGAGCTTTTCCACAAAGAGATTTCCAAATTAACAACCTGCCAG 3000
QY 3001 TCTGAGAAATCTCAAAACATCCGACGATCTGAGGCGGGCTTGGGATGGGAC 3060
Db 3001 TCTGAGAAATCTCAAAACATCCGACGATCTGAGGCGGGCTTGGGATGGGAC 3060
QY 3061 TGCCGCGCGGGTCTCTGAAAGAGATGCTGCGCGAGGACACACACACAGCAGCGCTG 3120
Db 3061 TGCCGCGCGGGTCTCTGAAAGAGATGCTGCGCGAGGACACACACACAGCAGCGCTG 3120
QY 3121 TGTGTGCGGCGGAGTCCGCTGCGGTCGCGGTGAGCAGCGGCTGTGTGTGGCGGGG 3180
Db 3121 TGTGTGCGGCGGAGTCCGCTGCGGTCGCGGTGAGCAGCGGCTGTGTGTGGCGGGG 3180
QY 3181 AGAGCCATTGTTCTGAGCGCTTACCGAGCCCGCGCTCGCCGGGAGGAGCGGGCT 3240
Db 3181 AGAGCCATTGTTCTGAGCGCTTACCGAGCCCGCGCTCGCCGGGAGGAGCGGGCT 3240
QY 3241 TCCCGGCTCCCAAGCTCCAGATCTCTGGGTGAGTCCAGCTCTCTCTGCGCAGCGGCTG 3300
Db 3241 TCCCGGCTCCCAAGCTCCAGATCTCTGGGTGAGTCCAGCTCTCTCTGCGCAGCGGCTG 3300
QY 3301 GGGGACCGGAAGACGGGACGGAGATTTAGTGTGGGCGGCGGCGGAGGTTTCCACT 3360
Db 3301 GGGGACCGGAAGACGGGACGGAGATTTAGTGTGGGCGGCGGCGGAGGTTTCCACT 3360
QY 3361 GTTCTCTGAGAACTTCCCGAGTCCGACACACCGCTTCTCCGTGTCGCCGAGGCGGCT 3420
Db 3361 GTTCTCTGAGAACTTCCCGAGTCCGACACACCGCTTCTCCGTGTCGCCGAGGCGGCT 3420
QY 3421 CCTGGGCTAGGCTCCGCGCGGCGGAGGCTTCCAGCGGCTTCCAGCGGCTTCCAGAGAAA 3480
Db 3421 CCTGGGCTAGGCTCCGCGCGGCGGAGGCTTCCAGCGGCTTCCAGCGGCTTCCAGAGAAA 3480
```

```
QY 3481 GCTCCCGACCGCGGATGCTCCGGGACAGAGCCAGCGGCGGTGGAAGAGAGCTGAGAAG 3540
Db 3481 GCTCCCGACCGCGGATGCTCCGGGACAGAGCCAGCGGCGGTGGAAGAGAGCTGAGAAG 3540
QY 3541 AGAAACAGAGGAGGAGGAGCGAGGAGCTGCGGCGAGAGGGAACACAGATTTCGCGCGA 3600
Db 3541 AGAAACAGAGGAGGAGGAGCGAGGAGCTGCGGCGAGAGGGAACACAGATTTCGCGCGA 3600
QY 3601 GCCAATGGCAACCGCAGGACGAGGTGGCACCAAAATTCCTTCCGCGCAATGACGAGCCGA 3660
Db 3601 GCCAATGGCAACCGCAGGACGAGGTGGCACCAAAATTCCTTCCGCGCAATGACGAGCCGA 3660
QY 3661 GTTTACAGAAAGCTCATTTAGCATTTCCCAAGAGCGAGGGGACAGAGCGGCTGG 3720
Db 3661 GTTTACAGAAAGCTCATTTAGCATTTCCCAAGAGCGAGGGGACAGAGCGGCTGG 3720
QY 3721 TGTGTGTCTGCTGCTGCGGACATCCCGCGGCTGCTGCGGTGCGCGGAGCTCGGCC 3780
Db 3721 TGTGTGTCTGCTGCTGCGGACATCCCGCGGCTGCTGCGGTGCGCGGAGCTCGGCC 3780
QY 3781 TCTGTCTCTCTCCCTCCCTCCCTTACCTCCAGCGGAGCCCGCGGCGCAGTCAACTCC 3840
Db 3781 TCTGTCTCTCTCCCTCCCTCCCTTACCTCCAGCGGAGCCCGCGGCGCAGTCAACTCC 3840
QY 3841 TCGCACTTTGCGGCTGCTGCGGAGGATTAAGAGGGGCTGAGGAAATACCGGACCGGT 3900
Db 3841 TCGCACTTTGCGGCTGCTGCGGAGGATTAAGAGGGGCTGAGGAAATACCGGACCGGT 3900
QY 3901 CACCGGTGCGAGCTTAGCTTTTAAATTCGCGGCTCGGGACCTCCAGCACCGGG 3958
Db 3901 CACCGGTGCGAGCTTAGCTTTTAAATTCGCGGCTCGGGACCTCCAGCACCGGG 3958
```

RESULT 2

```
US-10-087-192-10
; Sequence 10, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 37698
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-10
```

```
Query Match 99.3%; Score 3929.8; DB 13; Length 37698;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3953; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 3 GAATTCATCCATTTAAATCATACAAATTTAATGCTTTTAGTATATTTCACAGGTTTGCAT 62
Db 6280 GAATTCATCCATTTAAATCATACAAATTTAATGCTTTTAGTATATTTCACAGGTTTGCAT 6339
QY 63 CCATCAATCCCAATTTAGAACAGTTTATTACTCCAAAATAAACCCTCAATTCCTTAG 122
Db 6340 CCATCAATCCCAATTTAGAACAGTTTATTACTCCAAAATAAACCCTCAATTCCTTAG 6399
QY 123 CCATCAACCCCAACATCTCTCCATCTCTTCCAGCCCTGGGCAACCACTTACTT 182
Db 6400 CCATCAACCCCAACATCTCTCCATCTCTTCCAGCCCTGGGCAACCACTTACTT 6459
```

QY 183 TCTGTCCTATAAATTTGCCAATTTCTGGACATTTTCATATAAATGAAGCAAAACAATGT 242
Db 6460 TCTGTCCTATAAATTTGCCAATTTCTGGACATTTTCATATAAATGAAGCAAAACAATGT 6519
QY 243 GAGACTTTGTGACTGGCTGCTTTCATCTAGCAATTCATATTTTAAAGGCTCATATGTATACA 302
Db 6520 GAGACTTTGTGACTGGCTGCTTTCATCTAGCAATTCATATTTTAAAGGCTCATATGTATACA 6579
QY 303 GTACTTAGCAGTACATTCATTTCTTTTATTTCTCAATGGTATTCACATGTTGGGTATCC 362
Db 6580 GTACTTAGCAGTACATTCATTTCTTTTATTTCTCAATGGTATTCACATGTTGGGTATCC 6639
QY 363 CATATCATATATTATAGACAGATTTCTCATCTGTCAACCCAGGCTGGAGTGCAGTGGCAC 422
Db 6640 CATATCATATATTATAGACAGATTTCTCATCTGTCAACCCAGGCTGGAGTGCAGTGGCAC 6699
QY 423 AATCATAGCTCACTGTAACTCAAACTCTGGCTCAAGTGAATCTACTACCTCAGCCTC 482
Db 6700 AATCATAGCTCACTGTAACTCAAACTCTGGCTCAAGTGAATCTACTACCTCAGCCTC 6759
QY 483 CAGAGTAGCTAGGACTACAGGCACACACAGCCATACCTGGCTAAATTTTAAATTT 542
Db 6760 CAGAGTAGCTAGGACTACAGGCACACACAGCCATACCTGGCTAAATTTTAAATTT 6819
QY 543 TCATTTTATGTAATTCATTTTCTTTTGTGTTGTTGTTGTTGATAGGCTCTCAC 602
Db 6820 TCATTTTATGTAATTCATTTTCTTTTGTGTTGTTGTTGTTGATAGGCTCTCAC 6879
QY 603 TTTGTTATCCAGGCTGGAGGAGTGGCATGGTGAACAGCTGAGCAGCTTACCTTCCTGG 662
Db 6880 TTTGTTATCCAGGCTGGAGGAGTGGCATGGTGAACAGCTGAGCAGCTTACCTTCCTGG 6939
QY 663 GCTCAAGTGATCCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGCTGTACC 722
Db 6940 GCTCAAGTGATCCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGCTGTACC 6999
QY 723 ATGCTGCTGATATTTTCTTTTCTTTGAAACAGGATACACTGTGTTGCCAGGCTGGAGT 782
Db 7000 ATGCTGCTGATATTTTCTTTTCTTTGAAACAGGATACACTGTGTTGCCAGGCTGGAGT 7059
QY 783 ACAGTGGCGTAAATATAGCTCACTCAGCCTCCCTCTGGCTCAAGCAATCCGCTGGC 842
Db 7060 ACAGTGGCGTAAATATAGCTCACTCAGCCTCCCTCTGGCTCAAGCAATCCGCTGGC 7119
QY 843 CTCAGCATCCTGAGTAGCTGGGACTACAGGCTTGTGCCACAGGCCAGCTAAGTTTAA 902
Db 7120 CTCAGCATCCTGAGTAGCTGGGACTACAGGCTTGTGCCACAGGCCAGCTAAGTTTAA 7179
QY 903 AAAATGATTTTGTATAGAGAGGCTCTGTATGTTGCTCAGGCTGTATTTTATTTGTT 962
Db 7180 AAAATGATTTTGTATAGAGAGGCTCTGTATGTTGCTCAGGCTGTATTTTATTTGTT 7239
QY 963 GAGACAAGGCTCTCATATGTTGCCATGATCCCCACCTCCACTTCCCAAAAGTGTCTATC 1022
Db 7240 GAGACAAGGCTCTCATATGTTGCCATGATCCCCACCTCCACTTCCCAAAAGTGTCTATC 7299
QY 1023 TTATCTGTTCAATTAGTCAGTGTACAGACATTTAGGTTGTTTCCACTTTTGTACCAATTATG 1082
Db 7300 TTATCTGTTCAATTAGTCAGTGTACAGACATTTAGGTTGTTTCCACTTTTGTACCAATTATG 7359
QY 1083 AATAATACCTCAGTGAATATTCATGTATACATTTGTGTGGGCATATGTTTTCATTTCTGT 1142
Db 7360 AATAATACCTCAGTGAATATTCATGTATACATTTGTGTGGGCATATGTTTTCATTTCTGT 7419
QY 1143 TGGGTTTATATCTAGAGTGGAAATGCTGGATCCCGGCTAATATTTTTCACAGGAGAGTT 1202
Db 7420 TGGGTTTATATCTAGAGTGGAAATGCTGGATCCCGGCTAATATTTTTCACAGGAGAGTT 7479
QY 1203 CAGGGGAAGAAAACTTGGGAAAAATGAAGCATGTTTAGAAATCAGCAAGAGTGCAGGGGT 1262
Db 7480 CAGGGGAAGAAAACTTGGGAAAAATGAAGCATGTTTAGAAATCAGCAAGAGTGCAGGGGT 7539
QY 1263 TTTTCGGAGTTTATTTTATATTTCTGTGTGACAAATGTGACATTTGATGAAGATACAGATT 1322

Db 7540 TTTTCGGAGTTTATTTTATATTTCTGTTGACAAATGTGCAATTTGATGAAGATACAAGTT 7599
QY 1323 ATACTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1382
Db 7600 ATACTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7659
QY 1383 CACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATAATAAAAAATCACAATTT 1442
Db 7660 CACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATAATAAAAAATCACAATTT 7719
QY 1443 TTTTTTTTTTTTTTGAGAAAGAGCTTTGCTCTTTTCACTTGGCTGGAGGAGTGGTGTG 1502
Db 7720 TTTTTTTTTTTTTTGAGAAAGAGCTTTGCTCTTTTCACTTGGCTGGAGGAGTGGTGTG 7779
QY 1503 ATCTCAGCTCACTGCAACTTTGCTCTCCGGTTCAGCAATTTCTCCTGCTTCAGCCTCC 1562
Db 7780 ATCTCAGCTCACTGCAACTTTGCTCTCCGGTTCAGCAATTTCTCCTGCTTCAGCCTCC 7839
QY 1563 CAACTAGCTGGGACTACAGGCACTTCCCACTGCTTCAAGTGAATTTTGTATTTTAGTA 1622
Db 7840 CAACTAGCTGGGACTACAGGCACTTCCCACTGCTTCAAGTGAATTTTGTATTTTAGTA 7899
QY 1623 GAGATGGGANTTCACTTTGTTGGCAAGCTGGTCTCAAACTTTTGTGCTGCTCATATTTGTT 1682
Db 7900 GAGATGGGANTTCACTTTGTTGGCAAGCTGGTCTCAAACTTTTGTGCTGCTCATATTTGTT 7959
QY 1683 GTAACTATTTGTTTCTTGTCTGAGTGGGCTCCAGCAAAATAAATAAATCTTAGAA 1742
Db 7960 GTAACTATTTGTTTCTTGTCTGAGTGGGCTCCAGCAAAATAAATAAATCTTAGAA 8019
QY 1743 TCCAAATCAGTGTGTTGTTGTTGACCTGTCTCACTTGAGAACCAAGTGTGCCAGGGCT 1802
Db 8020 TCCAAATCAGTGTGTTGTTGTTGACCTGTCTCACTTGAGAACCAAGTGTGCCAGGGCT 8079
QY 1803 CAGAGTAGAGGTGATCTCTGCTCGAAGAGAAATAGAAATATTTCTCCGGGCGAG 1862
Db 8080 CAGAGTAGAGGTGATCTCTGCTCGAAGAGAAATAGAAATATTTCTCCGGGCGAG 8139
QY 1863 GCGTGTGGCTCANGCTTAAATCCAGCACTTTGGGAGGCCAAGGCAATGTGGATCACCT 1922
Db 8140 GCGTGTGGCTCANGCTTAAATCCAGCACTTTGGGAGGCCAAGGCAATGTGGATCACCT 8199
QY 1923 GAGTTCAGGAGTTCAAAACCAAGCCTGCGCAACATGGTGAACCCCGTCTCTACTAAAAAT 1982
Db 8200 GAGTTCAGGAGTTCAAAACCAAGCCTGCGCAACATGGTGAACCCCGTCTCTACTAAAAAT 8259
QY 1983 ACAAAAAATTAGCTTAAGTGGTGGCGCATGCCCTGTAAATCCAGCTACTTGGGAGGTGA 2042
Db 8260 ACAAAAAATTAGCTTAAGTGGTGGCGCATGCCCTGTAAATCCAGCTACTTGGGAGGTGA 8319
QY 2043 GGCAAGGAATTTCTTGAACCCGGGAGGAGAGGTTGCAAGTGAAGCGAGATCAACCACT 2102
Db 8320 GGCAAGGAATTTCTTGAACCCGGGAGGAGAGGTTGCAAGTGAAGCGAGATCAACCACT 8379
QY 2103 GCATCCAGCTGGGAGAGAGGAGACTTCTCTCAAAAAACAAAAACAAAAAAT 2162
Db 8380 GCATCCAGCTGGGAGAGAGGAGACTTCTCTCAAAAAACAAAAACAAAAAAT 8439
QY 2163 TAAGCAATTTAGACATTTGACAGAGAACTTGAAGGGGTGAGACCAAGTACAGATTCTG 2222
Db 8440 TAAGCAATTTAGACATTTGACAGAGAACTTGAAGGGGTGAGACCAAGTACAGATTCTG 8499
QY 2223 TGCCACATGCCAAGTACTTCTGAGGCAATGATGAGTGTGCTCAATCTGAAATCATC 2282
Db 8500 TGCCACATGCCAAGTACTTCTGAGGCAATGATGAGTGTGCTCAATCTGAAATCATC 8559
QY 2283 CAGTCTTTGTTTCAAGAACTTTTCAACCGGACAGGAGGCTGGAATGCAAGTCTCTCTG 2342
Db 8560 CAGTCTTTGTTTCAAGAACTTTTCAACCGGACAGGAGGCTGGAATGCAAGTCTCTCTG 8619
QY 2343 TCACITGGCAGAGAGTTGGGCTTGAACCTGAGACAGTGGCCAAAGAGGAGTCTTAG 2402

Db 8620 TCACTGCCAGAGAGTTGGCC-TGACCTTGAGACCACTGGCCAAAGAGAGCTGCTTAG 8678
 Qy 2403 TCTACCTCCAGGAAATCCAGAGTCTTCTCTCTCGGAGTCAATCATTTGGCGCAGCA 2462
 Db 8679 TCTACCTCCAGGAAATCCAGAGTCTTCTCTCTCGGAGTCAATCATTTGGCGCAGCA 8738
 Qy 2463 CTCGATTTTCTCTCTCTCCAGGGAAGAGTCTTAGGGCAGTATTTGGGAAAGACATG 2522
 Db 8739 CTCGATTTTCTCTCTCTCCAGGGAAGAGTCTTAGGGCAGTATTTGGGAAAGACATG 8798
 Qy 2523 GGCATGGAAGACACCGGGTGAATGATAGCTCTGCTGGTCTGAGTCTCATGGTAAGG 2582
 Db 8799 GGCATGGAAGACACCGGGTGAATGATAGCTCTGCTGGTCTGAGTCTCATGGTAAGG 8858
 Qy 2583 CTCCTACAGACACGGAAGATGGGGCAGACAGGACAGATCAGTAGGGTCAGACATCTC 2642
 Db 8859 CTCCTACAGACACGGAAGATGGGGCAGACAGGACAGATCAGTAGGGTCAGACATCTC 8918
 Qy 2643 AGGACCGAGGGCAATATGGTCTCTGAGCAGGGAATTAAGAGCTTGGGCTCTCATATGGTGT 2702
 Db 8919 AGGACCGAGGGCAATATGGTCTCTGAGCAGGGAATTAAGAGCTTGGGCTCTCATATGGTGT 8978
 Qy 2703 TTCTGGGCTCAACTGCCAGCTCCGCTACTTACTGGTCTGTGACCAATGAGGCAAGTTATT 2762
 Db 8979 TTCTGGGCTCAACTGCCAGCTCCATCACTTACTGGTCTGTGACCAATGAGGCAAGTTATT 9038
 Qy 2763 CCATCTCTCCATATCTCTTCTCTCACTTTAAATGGAATTAATGGGTACCCACTCCCA 2822
 Db 9039 CCATCTCTCCATATCTCTTCTCTCACTTTAAATGGAATTAATGGGTACCCACTCCCA 9098
 Qy 2823 GGGTCACAGAGAGCTTACAGAAACGATCTTGTGAAATGGCTTGCAGTAATTAATCAA 2882
 Db 9099 GGGTCACAGAGAGCTTACAGAAACGATCTTGTGAAATGGCTTGCAGTAATTAATCAA 9158
 Qy 2883 TACTGCGAGCTATCTTATTTCCATCCATCCAGCCCTTTGCGCTGCTGCTGGTGAAGACA 2942
 Db 9159 TACTGCGAGCTATCTTATTTCCATCCATCCAGCCCTTTGCGCTGCTGCTGGTGAAGACA 9218
 Qy 2943 CATGTGAGTGTCTTCTGACGGTTTCCAAAGAGATTCAAAATTAACCTGCCAGTC 3002
 Db 9219 CATGTGAGTGTCTTCTGACGGTTTCCAAAGAGATTCAAAATTAACCTGCCAGTC 9278
 Qy 3003 TGAAGATCTCCAAACATCCCGACGATCCTGGAGCGCGGGCTTGGGATGGGACTG 3062
 Db 9279 TGAAGATCTCCAAACATCCCGACGATCCTGGAGCGCGGGCTTGGGATGGGACTG 9338
 Qy 3063 CCGCCCGGGTCTGAAACAGATCGGTGCGGCAAGGACACACACACAGCCCTGTG 3122
 Db 9339 CCGCCCGGGTCTGAAACAGATCGGTGCGGCAAGGACACACACACAGCCCTGTG 9398
 Qy 3123 TGTGCGGCGGAGTCCGGTGGGTGCGGTGAGCAGCGGTGCTGGTGGGCGGGCAG 3182
 Db 9399 TGTGCGGCGGAGTCCGGTGGGTGCGGTGAGCAGCGGTGCTGGTGGGCGGGCAG 9458
 Qy 3183 AGCATTTGTCGAGGCTACCGAGCCCGCGCTCGCCGGGAGGAGCGGGGCTTC 3242
 Db 9459 AGCATTTGTCGAGGCTACCGAGTCCCGCGCTCGCCGGGAGGAGCGGGGCTTC 9518
 Qy 3243 CCGGCT-CCCCAGCTCAGATCTCTGGGTGGGTGCGCAGCTCTCTCTGCGACCGGCTTG 3301
 Db 9519 CCGGCTCCCCAGCTCAGATCTCTGGGTGGGTGCGCAGCTCTCTCTGCGACCGGCTTG 9578
 Qy 3302 GGGACGGGAAGACGGACGGAGATGTTAGTGGTGGGCGCCCGCCCGAGGGTTACCACTG 3361
 Db 9579 GGGACGGGAAGACGGACGGAGATGTTAGTGGTGGGCGCCCGCCCGAGGGTTACCACTG 9638
 Qy 3362 TTTCCTGAGAACTTCCCACTGCGCCCAACCCAGCTTCTCGGTGTCGCCGAGGCGCGGTC 3421
 Db 9639 TTTCCTGAGAACTTCCCACTGCGCCCAACCCAGCTTCTCGGTGTCGCCGAGGCGCGGTC 9698
 Qy 3422 CTGGGCTAGGCTCGCGGCCCGACCCCAACCCGGGTCCCGAGGCCCTTCCAGAGAGAAAG 3481
 Db 9699 CTGGGCTAGGCTCGCGGCCCGACCCCAACCCGGGTCCCGAGGCCCTTCCAGAGAGAAAG 9758

Qy 3482 CTCCCGACGGGATCCCGGCGAGAGCCCGAGCGGCGGCTGGAGAGAGCTGAGAGGA 3541
 Db 9759 CTCCCGACGGGATCCCGGCGAGAGCCCGAGCGGCGGCTGGAGAGAGCTGAGAGGA 9818
 Qy 3542 GAAACAGAGGGAGGGGAGCTGGGAGCTGGCGGAGAGGAGGAAACAGCAGATTGCGCCGAG 3601
 Db 9819 GAAACAGAGGGAGGGGAGCTGGGAGCTGGCGGAGAGGAGGAAACAGCAGATTGCGCCGAG 9878
 Qy 3602 CCAATGCAACGCGAGACAGAGTGGCAACAAATTCCTTCGGCAATGACGAGCCCGAG 3661
 Db 9879 CCAATGCAACGCGAGACAGAGTGGCAACAAATTCCTTCGGCAATGACGAGCCCGAG 9938
 Qy 3662 TTTTACAGAACCTCATTTAGCATTTCCCGAGGCGAGGGGCGAGAGCCCGGGTGGT 3721
 Db 9939 TTTTACAGAACCTCATTTAGCATTTCCCGAGGCGAGGGGCGAGAGCCCGGGTGGT 9998
 Qy 3722 GTGCTGTCGCTGTCGCGAGCATCCCGCGCGCTGCTGCGGTGCGCGAGCCCTCGGCT 3781
 Db 9999 GTGCTGTCGCTGTCGCGAGCATCCCGCGCGCTGCTGCGGTGCGCGAGCCCTCGGCT 10058
 Qy 3782 CTGCTCTCTCCCTCCCGGCTTACTCTCACGCGGAGCCCGCGCGCTGACTCTCT 3841
 Db 10059 CTGCTCTCTCCCTCCCGGCTTACTCTCACGCGGAGCCCGCGCGCTGACTCTCT 10118
 Qy 3842 CGCACTTTGCGCTGCTTGGCAGCGATAAAGGGGCTGAGGAAATACCGGACACCGTC 3901
 Db 10119 CGCACTTTGCGCTGCTTGGCAGCGATAAAGGGGCTGAGGAAATACCGGACACCGTC 10178
 Qy 3902 ACCGTTGCCAGCTCTAGCCCTTTAAATTTCCCGCTCGGGACCTCCACGACCGCGG 3958
 Db 10179 ACCGTTGCCAGCTCTAGCCCTTTAAATTTCCCGCTCGGGACCTCCACGACCGCGG 10235

RESULT 3

US-10-027-632-24494/c
 ; Sequence 24494, Application US/10027632
 ; Publication No. US20020198371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE OF INVENTION: Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 24494
 ; LENGTH: 631
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-10-027-632-24494

Query Match 15.9%; Score 630.6; DB 13; Length 631;
 Best Local Similarity 99.8%; Pred. No. 1.7e-158;
 Matches 630; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 968 AGGCTCTCATATGTTGGCATGATCCCGCCACCTCCACCAAGTGTCTATCTATC 1027


```

Db 631 AAGTCTCACTATGTTGCCATGATCCCCACCTCCACCTCCCAAAAGTGTCTCATCTTATC 572
QY 1028 TGTTCATTAGTCTAGTTGACAGACATTTAGGTTGTTTCCACCTTTTGGACCAATTATGAATAA 1087
Db 571 TGTTCATTAGTCTAGTTGACAGACATTTAGGTTGTTTCCACCTTTTGGACCAATTATGAATAA 512
QY 1088 TACTCCAGTGAATATTCATGATATACATTTGTTGGGCAATGTTTTCATTTCTGTTGGGT 1147
Db 511 TACTCCAGTGAATATTCATGATATACATTTGTTGGGCAATGTTTTCATTTCTGTTGGGT 452
QY 1148 TTATATCTAGAGTGAATTTGCTGGATCCCGGTTAATATTTTGCACAGGAGAGTTCCAGGG 1207
Db 451 TTATATCTAGAGTGAATTTGCTGGATCCCGGTTAATATTTTGCACAGGAGAGTTCCAGGG 392
QY 1208 GAAGAAAACTTGGGAAAAATGAAGCATGTTTGAAGAAATCAGCAAGAGTGCAGGGGTTTTTC 1267
Db 391 GAAGAAAACTTGGGAAAAATGAAGCATGTTTGAAGAAATCAGCAAGAGTGCAGGGGTTTTTC 332
QY 1268 GGAGTTTATTTTATATTTCTGTTGACAAATGTCAGTTTGCAGTTTGCATGAAGATCAAGTTTACT 1327
Db 331 GGAGTTTATTTTATATTTCTGTTGACAAATGTCAGTTTGCAGTTTGCATGAAGATCAAGTTTACT 272
QY 1328 AAGTGAAGTGAATTAAGGCTGAATAGGCGTTTCAGAGTAAATCATGAAGCACCT 1387
Db 271 AAGTGAAGTGAATTAAGGCTGAATAGGCGTTTCAGAGTAAATCATGAAGCACCT 212
QY 1388 TGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATATAAAATACAAATTTTTT 1447
Db 211 TGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATATAAAATACAAATTTTTT 152
QY 1448 TTTTTTTTTGAAGAGAGTCTTCTTTTACCTCGCTTGAAGGAGAGTGGTGTGATCTC 1507
Db 151 TTTTTTTTTGAAGAGAGTCTTCTTTTACCTCGCTTGAAGGAGAGTGGTGTGATCTC 92
QY 1508 AGCTCACTCAACTTTTCGCTCCCGGTTCAAGCAATTTCTCTGCTTCAAGCCTTCCCAAGT 1567
Db 91 AGCTCACTCAACTTTTCGCTCCCGGTTCAAGCAATTTCTCTGCTTCAAGCCTTCCCAAGT 32
QY 1568 AGCTGGGACTACAGGACATTTCCACCATGCC 1598
Db 31 AGCTGGGACTACAGGACATTTCCACCATGCC 1

```

```

RESULT 4
US-10-027-632-24494/c
; Sequence 24494, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24494
; LENGTH: 631
; TYPE: DNA

```

```

; ORGANISM: Human
; US-10-027-632-24494
Query Match 15.9%; Score 630.6; DB 16; Length 631;
Best Local Similarity 99.8%; Pred. No. 1.7e-158; Indels 0; Gaps 0;
Matches 630; Conservative 1; Mismatches 0;
QY 968 AAGTCTCACTATGTTGCCATGATCCCCACCTCCACCTTCCCAAAAGTGTCTCATCTTATC 1027
Db 631 AAGTCTCACTATGTTGCCATGATCCCCACCTCCACCTTCCCAAAAGTGTCTCATCTTATC 572
QY 1028 TGTTCATTAGTCTAGTTGACAGACATTTAGGTTGTTTCCACCTTTTGGACCAATTATGAATAA 1087
Db 571 TGTTCATTAGTCTAGTTGACAGACATTTAGGTTGTTTCCACCTTTTGGACCAATTATGAATAA 512
QY 1088 TACTCCAGTGAATATTCATGATATACATTTGTTGGGCAATGTTTTCATTTCTGTTGGGT 1147
Db 511 TACTCCAGTGAATATTCATGATATACATTTGTTGGGCAATGTTTTCATTTCTGTTGGGT 452
QY 1148 TTATATCTAGAGTGAATTTGCTGGATCCCGGTTAATATTTTGCACAGGAGAGTTCCAGGG 1207
Db 451 TTATATCTAGAGTGAATTTGCTGGATCCCGGTTAATATTTTGCACAGGAGAGTTCCAGGG 392
QY 1208 GAAGAAAACTTGGGAAAAATGAAGCATGTTTGAAGAAATCAGCAAGAGTGCAGGGGTTTTTC 1267
Db 391 GAAGAAAACTTGGGAAAAATGAAGCATGTTTGAAGAAATCAGCAAGAGTGCAGGGGTTTTTC 332
QY 1268 GGAGTTTATTTTATATTTCTGTTGACAAATGTCAGTTTGCAGTTTGCATGAAGATCAAGTTTACT 1327
Db 331 GGAGTTTATTTTATATTTCTGTTGACAAATGTCAGTTTGCAGTTTGCATGAAGATCAAGTTTACT 272
QY 1328 AAGTGAAGTGAATTAAGGCTGAATAGGCGTTTCAGAGTAAATCATGAAGCACCT 1387
Db 271 AAGTGAAGTGAATTAAGGCTGAATAGGCGTTTCAGAGTAAATCATGAAGCACCT 212
QY 1388 TGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATATAAAATACAAATTTTTT 1447
Db 211 TGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATATAAAATACAAATTTTTT 152
QY 1448 TTTTTTTTTGAAGAGAGTCTTCTTTTACCTCGCTTGAAGGAGAGTGGTGTGATCTC 1507
Db 151 TTTTTTTTTGAAGAGAGTCTTCTTTTACCTCGCTTGAAGGAGAGTGGTGTGATCTC 92
QY 1508 AGCTCACTCAACTTTTCGCTCCCGGTTCAAGCAATTTCTCTGCTTCAAGCCTTCCCAAGT 1567
Db 91 AGCTCACTCAACTTTTCGCTCCCGGTTCAAGCAATTTCTCTGCTTCAAGCCTTCCCAAGT 32
QY 1568 AGCTGGGACTACAGGACATTTCCACCATGCC 1598
Db 31 AGCTGGGACTACAGGACATTTCCACCATGCC 1

```

```

RESULT 5
US-09-792-468B-1
; Sequence 1, Application US/09792468B
; Publication No. US20030157552A1
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Brownlie, Alison J.
; APPLICANT: Ntambi, James M.
; APPLICANT: Miyazaki, Makoto
; APPLICANT: Gray-Keller, Mark P.
; APPLICANT: Attie, Alan D.
; TITLE OF INVENTION: Methods and Compositions Using Stearoyl-CoA Desaturase
; TO IDENTIFY TRIGLYCERIDE REDUCING THERAPEUTIC AGENTS
; FILE REFERENCE: 760050-7
; CURRENT APPLICATION NUMBER: US/09/792,468B
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: U.S. 60/184,526
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: U.S. 60/221,697
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: U.S. 60/255,771

```

```
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-792-468B-1

Query Match      12.6%; Score 498.4; DB 10; Length 617;
Best Local Similarity 99.8%; Pred. No. 5.7e-123;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3459 CCACAGCCCTTCCAGAGAGAAAGTCCCGACGCGGGATCCGGGACAGAGGCCACAGCGGCG 3518
Db 4 CCCC GCCCTTCCAGAGAGAAAGTCCCGACGCGGGATCCGGGACAGAGGCCACAGCGGCG 63

Qy 3519 GGTGAAGAGAGCTGAGAGAGAGAAACAGAGGGGAGGGAGCGAGAGCTGGCGGAG 3578
Db 64 GGTGAAGAGAGCTGAGAGAGAGAAACAGAGGGGAGGGAGCGAGAGCTGGCGGAG 123

Qy 3579 AGGGAACAGAGATTGCGCGAGCCAAATGGCAACGCGAGGACGAGGTGGCAACAAATCC 3638
Db 124 AGGGAACAGAGATTGCGCGAGCCAAATGGCAACGCGAGGACGAGGTGGCAACAAATCC 183

Qy 3639 CTTCCGGCCAATGACGAGCCGGAGTTTACAGAAAGCTCATTTAGCAATTTCCCCAGAGGAGG 3698
Db 184 CTTCCGGCCAATGACGAGCCGGAGTTTACAGAAAGCTCATTTAGCAATTTCCCCAGAGGAGG 243

Qy 3699 GGCAGGGGAGAGCCGGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 3758
Db 244 GGCAGGGGAGAGCCGGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 303

Qy 3759 GCGGTGCGCGGAGCTCGGCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3818
Db 304 GCGGTGCGCGGAGCTCGGCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 363

Qy 3819 ACCGCGCGCGCAGTCAACTCTCGCACTTTGCGCCCTGCTTTGGCAGCGGATATAAGGGGG 3878
Db 364 ACCGCGCGCGCAGTCAACTCTCGCACTTTGCGCCCTGCTTTGGCAGCGGATATAAGGGGG 423

Qy 3879 CTGAGGAAATACCGACACAGGTGACCCGTTGCCAGCTCTAGCCTTTAAATTCGCGCTCG 3938
Db 424 CTGAGGAAATACCGACACAGGTGACCCGTTGCCAGCTCTAGCCTTTAAATTCGCGCTCG 483

Qy 3939 GGGACCTCCACGACCGCGG 3958
Db 484 GGGACCTCCACGACCGCGG 503

RESULT 6
US-10-087-192-814/c
; Sequence 814, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 814
; LENGTH: 59725
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (1)...(59725)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-814

Query Match      11.6%; Score 460; DB 13; Length 59725;
Best Local Similarity 58.9%; Pred. No. 2.7e-111;
Matches 1107; Conservative 0; Mismatches 660; Indels 111; Gaps 14;

Qy 365 TATCATATATTAGACAGAGTCTCTCACTCTGTCCACCCAGGCTGAGTGAGTGCGCAAA 424
Db 33012 TTTTCTTCTTTTGGACAGGGTCTTCTCTGTCCACCCAGGCTGGAAACCCAGTGCGCAAA 32953

Qy 425 TCATAGTCTCACTGTAAACCTCTCTGGGCTCAAGTGATCTTACTACTACCTTCAGCCTCCA 484
Db 32952 TCTCAGTCTCACTAAACCTCTCTGGCTCTCTGGTTCAGCAATTTCTCTCTCGCTCAGCTCT 32893

Qy 485 GAGTAGCTAGACATACAGGACACACAGCCATACCTGGCTAAATTTTTTTTTTTTAAATTTTC 544
Db 32892 GAGTAGCTAGACATACAGGAGGTGTGCCACACACAGATTAATTTTTTATATTTTATA 32833

Qy 545 ATTTTATGTATTCATTTCT- - - - -TTCTTTTTTTTGTGTGTGTGTGTGTGTGTGTGTGT 595
Db 32832 GAGTAGGGTTCATATGTTGGCCAGGCTGGCTTTTCTTTGTTGTTTTTTTGTGTGTGTGTGT 32773

Qy 596 GTCTCACTTTTACCAGGCTGGAGGGCAGTGGCATGTGTGACAGCTGA - - - - -GCAAGCCTTG 653
Db 32772 GTCTTACTCTTTTCCAGGCTAGAGTGCAATGCAATCATGCTCACTGCGAGCTCA 32713

Qy 654 ACTTCTCTGGGCTCAAGTGATCTCTCTCTCAGCTCCGAGCTCCGAGTAGCTGGGACTCAAAACA 713
Db 32712 ACTTCTTAGGCTCAAGCAATCTCTCCACCTCGGCTTCCCAAAGTGTCTGGAATTTACAGACA 32653

Qy 714 CGTGTCAACATGC - - - - -CTGGCTGATATTTTTTTTCTTGAACAGGGTATCACTCTGTGC 770
Db 32652 TGAGTCAACATCTTTTTTTTTTTTTTTTTTTTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 32593

Qy 771 CCAGGCTGGAGTACAGTGGGCTAATAATAGTCACTCAAGCTCCCTCTCTCTCTCTCTCTCTCTCT 830
Db 32592 CCAGGCTGGATGAGTGGGACCAACCTCGGCTCACTGCAACCTCTGCTCTCTCTCTCTCTCTCT 32533

Qy 831 CAATCCGCTGGCTCAGCATCTGAGTAGCTGGGACTGAGGCTTGTGCGACAGGCGGCA 890
Db 32532 CAATCTCTCGCTCAGCATCTGAGTAGCTGGGACTGAGGCGGATGCGCACCAATGCGCTG 32473

Qy 891 GCTAAGTTTTTAAAAAATGATTTTTTGTGTATAGAGAGGTCTGTCTATGTTGTCTCAGGCTGT 950
Db 32472 GCTAATTTTT - - - - -GTATTTTTTAGTAGATGGGGTTTCAGCATGTTGGCCAGACTGG 32419

Qy 951 ATTTTATTTGTGAGACAAGTCTCACTATGTTGCCATGATCCCGCCACCTCTCTCTCTCTCTCT 1010
Db 32418 TCTCAAACTCTGA - - - - -CCTCAATGATCTGCGCGCTCTGCGCTCTCTCTCTCTCTCT 32375

Qy 1011 AAGTGTCTATCTTATCTGTCTTATTAGTCACTGACAGACATTTAGGTTGTTTCCACTTT 1070
Db 32374 AAGTGTCTGGAATTTACAGGATGAGGACCTGCAACCGGCGGCTATTTCTTATTTTCACTTA 32315

Qy 1071 TTGACCATTTATGAATATATCTCCAGTGAATTTTCATGTATACATTTTGTGTGGGCTATGT 1130
Db 32314 CAGATGAGGGAACCAAGACCTCGAGAGTGAAGTCACTTTGCTTTGTATGACACAGGAGT 32255

Qy 1131 TTTTCAATTTCTGTGGTTTATATCTAGGAGTGAATTTGTGGATCCCGGGTAAATTTTGT 1190
Db 32254 CAGCCAGGGCTGGGATTCGTGCCAGGTAGGCTGTGCTCTCACAATTTTCTTTCTTTCTTT 32195

Qy 1191 ACAGGAGAGTTTACGGGGAAGAAAACCTTGGGAAATGAGAGCATGTTTAGAATATCAGCAA 1250
Db 32194 TCACCTTTTTTTTTTTTCTGAGATGGAGTTTCACTCTT - - - - -CCTGCCAGGCTGGAGTGC 32138

Qy 1251 GAGTGCAGGGGTTTTTCGGAGTTTTTATTTTATATCTGTTGACAAATGTGCAGTTTGTATG 1310
Db 32137 GAGACCACTCTATCAACATGGTGAACCCCAATGGTGGATCTCTAGCTCACTGTTTCTT 32078
```


Db 693 CTCAAGTATCTCTCGCTCAGCTGCTTAATAGCTAGAGGACAGGTTGTGCCACA 752
QY 885 GGCCAGCTAAGTTTAAAAAATGTTTTGGTATAGAGAGGCTCTTGCTATGTGCTCA 944
Db 753 TGCTAGCTAATTTTAAATT---TTTTTTGTAGAGATGGATCTCACTATGTGGCTA 808
QY 945 GGCTGTATTTTATTTGTAGACAGAGTCTCACTATGTTGCCATGATCCCCACCTCCA 1004
Db 809 GGCTGCTTCTGAACCTCGGGT-----TCAAGCAGTCTCTTTGGCTTGGCTCTCCGA 861
QY 1005 CTTCCCAAGTGCATCTTATCTGTTCAITPAGTCAGTTGACAGACATTTAGTTGTTTC 1064
Db 862 GTGTTGAGTTGGGGCGAAGCACTGCTCCAACTGCTCTTTTGTGTGATGTAG 921
QY 1065 CACTTTTGGACATTAATGAATAAATPACTCCAGTGAATATCATGTATACATTTGTGTGGC 1124
Db 922 TAACATATTGATGTTTTTACTTAGATCCATAAATCAATTAAGAGCTATACAATGATGATA 981
QY 1125 ATATGTTTTCATTTCTGTTGGTTTATATCTAGAGTGGAAATGCTGGATCCCGGTAAT 1184
Db 982 TTCTAGTCTTATTTCTTTCAITTAATAGCTGAATPACTTA-----TAAA 1029
QY 1185 ATTTTGACAGGAGTTCAGGGGAAGAAAACCTTGGGAAAATCAAGCATGTTAGAAAT 1244
Db 1030 GAGAAACCTCCCTTTCTAATCTTCGATTTACCCAGTGATATAATATTTGGAAGT 1089
QY 1245 CAGCAAGAGTCAGGGGTTTTTCGGAGTTTTTATTTTATATCTGTTGACAAAATGTCAGT 1304
Db 1090 CAGGATGAATATTTGTTCTTTCCCTGTTTACCAATTTTCAAGATCTGATTTGGTTC 1149
QY 1305 TTGATGAAGTACAAGTTATPACATGAGTGAAGTGAATTAAGGCTGGAATAGGGGTT 1364
Db 1150 ACTAGCATTTCTCAA--TTAGGCATTTGTTTTGTGGGGTTTTTTTTGTTATCATAACTC 1208
QY 1365 CAGGTAATAATCATGAAGCACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAACAAA 1424
Db 1209 TTGGGTTTAAACATTTTGAATGTTTCAATTCATTTGCATTTTTTTCCTTTCTAGTGA 1268
QY 1425 TAATAAAAA-----ATCACAATTTTTTTTTTTTTTTTGGAGAAAGTCTGCTCTTTC 1477
Db 1269 TTTTACTGATATATTTCTTCTTCTTCTTTTCTTTTGTGAGACGGAGTCTGCTCTGTT 1328
QY 1478 ACCCTGCTGGAGGAGGAGTGTGATCTCAGCTCACTGCAACTTTCGCCCTCCCGGGTTC 1537
Db 1329 GCGAGGCTGGAGTGCAGTGGTGCATCTCAGCTAACTGCAACTCGGCTCTAGGGTTC 1388
QY 1538 AAGCAATTTCTCTGCTTTCAGCTCCCAAGTACCTGAGTACAGGACATTCACCACATGC 1597
Db 1389 AAGTGATTTCTCTGCTCAGTCTCCTGAGTGTGGATTTACGCCATGACACACCTCGC 1448
QY 1598 CCAGCTGATTTTGTATTTTATAGTAGAGTGGATTTCACTTTTGTGGCAAGCTGGTCT 1657
Db 1449 CCAACT-AAATTTGTATTTTATAGATGGGTTTTCTCCATGTTGTCAGGCTGCTCT 1507
QY 1658 CAAACTTTT-----TGCTGTCAATATTTGTGTAATCTATTTGCTTGTGAGTAGGGC 1713
Db 1508 CGAACTCTTTGAOCTAGGTGATCTCCCGCTCGGCTCCCAAGTACTAGGATTAACAG 1567
QY 1714 CCGCAGACAAAATAAATCTTAGAATCCAAATCAGTGTCTGTTTGACCACTGTC 1773
Db 1568 CACCATGCCAGTGGCATATATATTTTCTTAATAAATCACTCAGCTTGATGATGAGGACGTG 1627
QY 1774 ACTTGAGAACCAAGTGTGACCAAGGCTCTAGAGTAGAGGTGATCTGCTCGAAAGAG 1833
Db 1628 TCTTAGACATCTTCTGTCCATGTTGCTAGTGCCTGGTCTTTCAATAAATGCAATGAA 1687
QY 1834 AATAGAAAT-----GAAATATTTCTCGGGCCAGGCGTGTGGCTCA 1875
Db 1688 TGAGTTGATATTTCTGACAGCATTTTAAAAATAGAGCTAGTGGCTGGGACGCTGCTCA 1747
QY 1876 TGCTCTAATCCAGCACTTTTGGAGGCAAGGCAATGATGATCACCTGAGGTGACAGGTT 1935
Db 1748 TGCTGTATCCAGCACTTTGGAGGCCAAGCGCGGATCAC--GAGGTGAGGAT 1805

QY 1936 CAAAACCAAGCTGGCCAAACATGTTGAAACCCGCTCTCTACTAAAAATACAAAAATTAGC 1995
Db 1806 -GAGACATTTCTGGCCACATGTTGAAACCCGCTCTCTACTAAAAATAC-AAAAATTAGC 1863
QY 1996 TAAAGTGTGGTGGCCATGCTGTAATCCAGCTACTTTGGAGGCTGAGGAGGAATTT 2055
Db 1864 CGGAGTGTGGTGGTGGCTTTAGTCCAGCTACTCGGAGGCTGAGGAGGAATGG 1923
QY 2056 CTTGAACCCGGAGGAGAGGTTGCACTGAAGCGAGATCACACCTGCACTCCAGCTG 2115
Db 1924 TATGAACCCGGAGGAGGAGCTTGCAGTGAAGTTCGCACTGCACTCCAGCTG 1983
QY 2116 GGGGAGAGAGCAGAGACTTCTCTCAAAAAAACAACAA 2156
Db 1984 GGGACAGAGTGAAGTCTCTCTCAAAAAAACAACAA 2024

RESULT 8

US-10-087-192-4/c
; Sequence 4, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FASLSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 23618
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(23618)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-4

Query Match 9.1%; Score 360.4; DB 13; Length 23618;
Best Local Similarity 69.3%; Pred. No. 8.3e-85;
Matches 503; Conservative 0; Mismatches 221; Indels 2; Gaps 2;

QY 1441 TTTTCTTTTGTGAGAAAGAGTCTTGTCTTTTTCACCTGGCTGGAGGCGAGTGGTG 1500
Db 11003 TTTCTTTTGTGAGACAGAGTCTGCTCTGTACCCAGGCTGGAATGAGTGGCG 10944
QY 1501 TGATCTCAGCTCAGTGCAACTTTGCGCTCCCGGTTTCAAGCAATTCCTCGTTCAGCCT 1560
Db 10943 CGATCTCGGCTCACTGCAACCTTCACTCATAGGTACACGCAATTCCTCCTCCCTCAGCCT 10884
QY 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCACATATGCCAGCTGATTTTGTATTTT 1620
Db 10883 CCCGAGTAGCTGGGACTACAGGCGGCCACCATGCTGGCTAAATTTTGTATTTT 10824
QY 1621 TAGAGATGGGATTTTCACTTTTGTGGCAAGTGGTCTCAAACTTTTGTCTCATTAATG 1680
Db 10823 TAGAGACGGGTTTCAACATATTTGGCCAGGCTGGTCTCGAATCTCTGACTTGTGATCG 10764
QY 1681 TTGTAACATATTTCTCTTTTGTCTGAGTAGGCCCCCAGACCAAAAAAATAAATCTTAG 1740
Db 10763 CCCACCTTGGCCTCCCAAGTCTGGGATTAACAGGCTGAGCCATGGCGCCAGCCTCAT 10704
QY 1741 AATCCAAATCAGTGTGTGGTT-TGACCACCTGCTACTTGAGAACCAACAGTGTGACGAGG 1799
Db 10703 TTTAGCTATTTCTTATTTTCTTATTTTCCATTTTATTTTCCATTTATGATCTACACACTGAGTTTATTA 10644

QY 1800 CCTCAGGAGTACAGGTGATCTCTGCTCGAAAGAGAAATAGAAATATTTCTCCGGGC 1859
DB 10643 CCTATGTAAGATGTGTTATATTTTCGNNNNNNNNNNNNNNNNNNNNNTATATTTTGGC 10584
QY 1860 CAGCGGTGGTCTCATGCTGTATATCCAGCAGCTTTGGAGGCCCAAGGCATGTGGATCA 1919
DB 10593 CAGGCATGGTGGTCTCATGCTGTATATCCAGCAGCTTTGGAGGGCTGAGGCGAGCGATCA 10524
QY 1920 CCTGAGGTACAGAGTTCAAAACACGCTGGCCCAACATGCTGAACCCCGCTCTCTACTAAA 1979
DB 10523 CCTGAGGTACAGAGTTGAGATCAGCTGGCCCAACATGCTGAACCCCGCTCTACTAAA 10464
QY 1980 AATCAAAAAATAGCTAAGTGTGGTGGCATGCTGTATATCCAGCTACTTTGGGAGGG 2039
DB 10463 AATAC-AAAAATTAACCGGGGCTGGTGGCAGCGCGCTGTATATCCAGGTACTTGGGAGGC 10405
QY 2040 TGAGCAGAGAAATTTCTTGAAACCGGAGGAGCAGAGGTTGCAGTGAAGCGAGATCACACC 2099
DB 10404 TGAGCAGAGAAATCGCTGTGTACCTGGAAGCAGAGGTTGCAGTGAAGCGAGACTGAACC 10345
QY 2100 ACTGCACCTCCAGCTCGGGGAGAGAGAGACTTCTCTCAAAAAACAAAAACAAAAAG 2159
DB 10344 ATTGCACCTCCAGCTGGGTGACAGAGCAAGACTCTGTGCGCAAAAAAATTAATA 10285
QY 2160 AATTAA 2165
DB 10284 ATTTTA 10279

RESULT 9
; Sequence 1584, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; PRIOR FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1584
; LENGTH: 36296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 Z82180
US-10-240-425-1584

Query Match 8.9%; Score 353.6; DB 13; Length 36296;
Best Local Similarity 57.2%; Pred. No. 7.5e-83;
Matches 956; Conservative 0; Mismatches 609; Indels 107; Gaps 13;

QY 522 GCTAATTTTTTTTTTAATTTTCAATTTATGTAATTCATTTCTTTCTTTTGTGTGT 581
DB 15559 GGTAATGTTGAGCTTCTCAAGCAGCAATTTATCTATTTATTTATTCATGTAATTT 15500
QY 582 TGTTTTGAATAGGTCTCACTTTGTTTACCAGGCTGGAGGCGAGTGGCATGGTGACAGC 641
DB 15499 TTTTATAGACAAGGCTCTCAATTCCTCACCAGGCTGGAGTGTGTGCAATCATAGCTC 15440
QY 642 TGAGCAGCCTTGACTTCTCTGGGCTCAAGTGATCTCTCGCTCGCCTCAGCCTCCCAAGTAGCTG 701

DB 15439 ACTCAGGCTCCAACTCTCTGGGCTCAAGTGATCTCAGCCTCCTGAGTAGTTG 15380
QY 702 GGACTACAAACACATGTCACCATGCTGGCTGATATTTTTTTTCTTGAACAGGGTATCA 761
DB 15379 GGACTGCAAGTGCATGCCCATGCTGGCTTTTTTTTTTTTTCAGACAGAGTTTCGCT 15320
QY 762 CTCTGTTGCCAGGCTGGAGTACAGTGGCGTAATAATAGCTCACTGCAAGCTCC-CTTCC 820
DB 15319 CT-TGTCACCCAAGCTGGAGTGCAGTGGCAGATCTTGGCTCACTGCAACCTCGCCTCC 15261
QY 821 TGGCTCAAGCAATCCGCTGGCTCAGCATCTCAGTAGCTGGAGCTACAGGCTTGTGCC 880
DB 15260 CAGGTTCAAGTGATTTCTCTGCTCAGCTCCTGAATAGCTGGGATTACAGGCAACACC 15201
QY 881 ACCAGGCCAGCTAAGTTTTAAAAATGATTTTTGTTATAGAGGAGGTCCTTGTCTATGTTG 940
DB 15200 ACCAGGCTGTCTAATTTTT-TGTATTTTAGTAGACAGGGTTTCCCATGTTG 15146
QY 941 CTCAGGCTGTATTTTTTATTTGTTGAGACAAAGTCTCACTATGTCCTCATGATCCCTCCACC 1000
DB 15145 GGCAAGCTGGTCTCTAACTCTGACCTCAGG-----TGATCCGCTGCC 15102
QY 1001 TCCACTTCCCAAGTGTCTCATTTATCTGTTTCATTAGTCAGTGCAGACATTTAGGTTG 1060
DB 15101 TCAGCCTCCCAAGTGTCTGGGATTACAGGTGTAGGGGCCATTT-TTTTTTTAAGGTAC 15044
QY 1061 TTTCCACTTTTGAACCAATTAATAATCTCCAGTGAATATTCATGCTATATACATTTGTGT 1120
DB 15043 AGATGGGCTCTTGCTATGTTGCTCAGGCTGGTCTCAAAATCTTTGGGCTCAAGTATCCTC 14984
QY 1121 GGGCATATGTTTTTCATTTCTGTTGGGTTTATATCTAGAGTGGAAATTGCTGGATCCCGGG 1180
DB 14983 CCTCCTCAGGCTCCCAAGTGTGGGATTATAGGAGGAGCCACCGACCTCAGG 14924
QY 1181 TAATATTTTACAGGCGAGGTTTCAAGGGGAGAAAACTTGGAAAAATGAACATGTTTAG 1240
DB 14923 TGACATTTTAATGGGAGAGGANGTGTGTGAGTGCATCCAAAGGGAGACGTCCTGATG 14864
QY 1241 AAATCAGCAAGTGCAGGGGTTTTTCGGAGTTTTATTTTATATTTCTGTGACAAAATGTG 1300
DB 14863 CTGTTAGAGCCAGCTAGAGTTTGGTCTTCCCTTAGTCTGAGGTAGGAGAGATTTTT 14804
QY 1301 CAGTTTGAAGAGATCAAGTTTATCTAAGTGAGAGTGAGAAATTAAGGCTGGAAATAGG 1360
DB 14803 CTGTGCCAAGAGCCCTGCAGTT-----14782
QY 1361 CGTTTCAGAGTAAATCATGAAGCACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAC 1420
DB 14781 -----CAGTGGTGGGACCTGACAACTCTCTTTATAGGAAGACACC 14741
QY 1421 AAAATAATAAAAAATCAAAATTTTTTTTTTTTTTTTGGAGAAAGAGTCTTGTCTTTCAAC 1480
DB 14740 CTCCTGTAGTCTTTTTTTTTTTTTTTTTTTTCTTTTTCAGACAAAGTCTTGTCTTTCA-C 14682
QY 1481 CTGCTGGAGGCGAGTGTGTGATCTCAGCTCACTGCAACTTTTGCCTCCCGGGTCTCAAG 1540
DB 14681 CAGGCTGGAGTGCAATGGCATGATCTCGGCTCACTGCAACCTCCACCTCTTGGGTTCAAG 14622
QY 1541 CAATTCCTGCTTCAGCTCCCAAGTAGCTGGAGCTACAGGCACTTCCCAACCATGCCCA 1600
DB 14621 CGATTTCTTGGCTTCAGCCTCCGAGTAGTGGGAGTACAGGCCACCGCCACCACTCTG 14562
QY 1601 GCTGATTTTTGTATTTTTTAGTAGATGGGATTTCACTTTTGGCCCAAGCTGCTCTCAA 1660
DB 14561 GCTAAATTTTGTATTTTATAGAGGGGAGCTTCACTGTTGGCCAGGATGGTCTCGA 14502
QY 1661 ACTTTTGTGTGATATTTGTGTAATATTTGTTTCC-----TTTTGCTGAGTAGGGC 1713
DB 14501 TCTCCTGACCTATGATCTGCTGCTCGGCTTTACAGGCATGAGCCACTGCGCCCGGCC 14442
QY 1714 CCCCAGACAAAAAATAAATCTTAGAATCCAAATCAGTGTGTTGGTTTGAACCATGTC 1773


```
Db 4155 ATAC-AAAAATTAGCCACGCTGTGGCGCATGCTATAGTCCCACTACTTTGAGAGGCT 4213
QY 2041 GAGGAGGAGAAATTTCTTGAACCCGGAGGAGGAGGTTGCAGTGAAGCCAGATCACACCA 2100
Db 4214 GAGGAGGAGAAATCACTTGAACCCAGGAGGAGGTTGCAGTGAAGCCAGATCACACCA 4273
QY 2101 CTGCACTCCAGCTGGGAGGAGGAGGAGTCTCTCTCAAAAAACAAAAACAAAAAGA 2160
Db 4274 CTGCACTCCAGCTGGGAGGAGGAGGAGTCTCTCTCAAAAAACAAAAACAAAAAGA 4333
QY 2161 ATTAAGCAAAATTAGACA 2177
Db 4334 ACAAAACAAAAACA 4350

RESULT 13
US-09-764-891-7809
; Sequence 7809, Application US/09764891
; Publication NO. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7809
; LENGTH: 29163
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7809

Query Match 8.8%; Score 347.4; DB 10; Length 29163;
Best Local Similarity 69.9%; Pred. No. 3e-81;
Matches 515; Conservative 0; Mismatches 211; Indels 11; Gaps 3;

QY 1441 TTTTCTTTTCTTTTGTGAAAGAGCTTGTCTTTTCACTTGTCCCTGGCTGGAGGCGAGTGTG 1500
Db 3623 TCTTTTCTTTTCTTTTCTGGAGTCTGTCTGTCCCGAGGCTGGAGTGCAGTGTG 3682
QY 1501 TGAATCAGCTCACTGCAACTTTTCCCTCCCGGTTCAAGCAATTTCTCTGCTTCAGCCT 1560
Db 3683 CGATCTCAGCTCACTGCAAACTCCCACTCCCGGTTCAAGCAATTTCTCTGCTTCAGCCT 3742
QY 1561 CCCAAGTAGCTGGGACTACAGCACTTCCCACTCCCGAGTGGTCTGCTGCTCATAATTG 1680
Db 3803 TACAGATGGGTTTCAACCATGTTGTCTAGGCTGGTCTCGAACTCTGACTTCATGATCCG 3862
QY 1681 TTGTAACATTTGTTCTTTTGTGAGGTAGGCGCCCAAGCAAAAAAATAATCTTAG 1740
Db 3863 TCCCTCCGACTCCCAAGTGTAGGATTACAGGCATGAGCACTGTGCCAGGCC---- 3918
QY 1741 AATCAAAATCAGTGTGTTGTGACCACTGTCACTTGAGAACACACAGTGTGACCAAGGC 1800
Db 3919 ---GCTAATTTATTTTGTACAGACAGATTTCTCCATGTTGCCAGGCTGATCTGGAA 3975
QY 1801 CTCAGGAGTAGAGTGTATCTCTGCTCGAAAGAGAAATAGAAATATTTCTCCGGGCC 1860
Db 3976 CTCCTGGGCGAAGTGAACCATCTGCTCGC---CTCCAAGTCTGGGATTACAGGCC 4032
QY 1861 AGGCGTGTGGCTATGCTGTATATCCAGCACTTTGGAGGCCCAAGCATGTGATCAC 1920
Db 4033 GGGTGGGTGGCTCACGCTGTGAATTCAGCACTTTGGAGGCCCAAGCATGGATCAG 4092
QY 1921 CTGAGGTGAGGTTCAAAACCCAGCTGGCCCAATGTTGAAACCCGCTCTCTACTAAA 1980
Db 4093 CTGAGGTCCGGAGTTCAAGACAGCCTGGCCCAAGAGGTGACACCTGTCTCTACTAAA 4152
```

```
QY 1981 ATACAAAAATTAGCTAAGTGTGGTGGCGCATGCTGTAATCCCACTACTTTGAGAGGCT 2040
Db 4153 ATAC-AAAAATAAGCCACGCTGTGGTGGCGCATGCTATAGTCCCACTACTTTGAGAGGCT 4211
QY 2041 GAGGAGGAGAAATTTCTTGAACCCGGAGGAGGAGTGTGCAGTGAAGCCAGATCACACCA 2100
Db 4212 GAGGAGGAGAAATCACTTGAACCCAGGAGGAGGAGTGTGCAGTGAAGCCAGATCACACCA 4271
QY 2101 CTGCACTCCAGCTGGGAGGAGGAGGAGTCTCTCTCAAAAAACAAAAACAAAAAGA 2160
Db 4272 CTGCACTCCAGCTGGGAGGAGGAGGAGTGTGCTCTCAAAACAAAAACAAAAACAAA 4331
QY 2161 ATTAAGCAAAATTAGACA 2177
Db 4332 ACAAAACAAAAACA 4348

RESULT 14
US-09-880-107-2278
; Sequence 2278, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2278
; LENGTH: 26928
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M17262
US-09-880-107-2278

Query Match 8.7%; Score 342.8; DB 9; Length 26928;
Best Local Similarity 57.5%; Pred. No. 4.9e-80;
Matches 848; Conservative 0; Mismatches 552; Indels 74; Gaps 10;

QY 735 TATTTTTTTTCTTGAACACAGGGTATCACTCTGTTGCCAGGCTGGAGTACAGTGGCGTAA 794
Db 3122 TTTTTTTTTTGGGAGTCAAGAGTCTTGTCTGTTGCCAGGCTGGAGTGCAGTGCACAGA 3181
QY 795 TAAATGCTCACTGAGCCTCC-CTCTGGGCTCAAGCAATCCGCTGGCTCCATCCT 853
Db 3182 TCTCAGCTCACTGAGCCTCCGCTCCCGGTTTCAAGCTATTCTCTGCTCTCAGCCTCC 3241
QY 854 GAGTAGCTGGGACTACAGGCTGTGCCACAGGCCAGCTTAAGTTTTAAAAAATGATTTT 913
Db 3242 AAATAGCTGAGCTATAGGCAAGCAGATCCATGCCAGCTAAATTTTT-----TTATTTT 3295
QY 914 TGGTATAGAGAGGCTTGTGCTATGCTGCTCAGGCTGTATTTTTATTGTTGAGACAAGTC 973
Db 3296 TAGTAGAGAGAGGCTGTCTCATGGTGGCCAGGTTGGTCTTGAAC----- 3341
QY 974 TCATATGTTGCCATGATCCCCCACTTCCAAAGTGTCTCATCTTATCTGTCA 1033
Db 3342 --CTGTCTCTCAAGTGTATCCACTGCTCCGCTCCCAAGTGTGGATTTGAGGATG 3399
QY 1034 TTAGTCACTGTGACAGACATTTAGTGTGTTTCCACTTTTTCACCATTTATGATATATCTCC 1093
Db 3400 AGACACCGCGCCGGCC-----TGCCTGTGCTCTTAAATGAGTTGTCCATTTGTAA 3454
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:19 ; Search time 14547.2 Seconds
(without alignments)
8124.892 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3958

Perfect score: 3958

Sequence: 1 aggaatcattccatttaaat.....gggacctccacgcgcgg 3958

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vri:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	649.4	16.4	1238	12	BM563325 AGENCOURT
2	491.4	12.4	579	12	BI559696 603252566
3	450.2	11.4	471	28	AQ760013 HS_3202_A
4	334	8.4	754	12	BI627092 603075763

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C	5	294.2	7.4	1910	11	BC035771	Homo sapi
	6	290.6	7.3	2076	11	BC035510	Homo sapi
C	7	290.6	7.3	2772	11	BC038630	Homo sapi
	8	283.4	7.2	2772	11	BC038630	Homo sapi
C	9	273.4	6.9	2076	11	BC035510	Homo sapi
	10	270	6.8	749	13	BU175345	AGENCOURT
	11	269.4	6.8	1172	12	BM546583	AGENCOURT
	12	269.4	6.8	1910	11	BC035771	Homo sapi
	13	259.4	6.6	929	9	AL580585	AL580585
	14	259.2	6.5	693	13	BU616025	BU616025
	15	255.8	6.5	2971	28	AF101960	AF101960
C	16	255.2	6.4	958	9	AL578531	AL578531
	17	255	6.4	694	9	AL596692	AL596692
	18	254.4	6.4	764	13	BU617735	BU617735
	19	252.4	6.4	940	13	EX424350	EX424350
	20	251	6.3	469	9	AL701463	AL701463
	21	251	6.3	560	9	AL708245	AL708245
	22	250.2	6.3	683	12	BI602109	BI602109
C	23	249.4	6.3	524	10	AW973992	AW973992
	24	249.2	6.3	580	12	BQ053726	AGENCOURT
	25	249	6.3	978	13	EX325356	EX325356
	26	248.4	6.3	1201	13	EX385602	EX385602
C	27	248.2	6.3	731	9	AI687343	AI687343
	28	247.2	6.2	775	12	BM905333	AGENCOURT
C	29	246.8	6.2	521	13	BU617227	BU617227
	30	246.8	6.2	598	12	BQ017808	UI-H-DF0-
C	31	246.8	6.2	616	13	BU617236	BU617236
C	32	246.2	6.2	508	28	AQ418545	AQ418545
	33	246.2	6.2	958	9	AL578531	AL578531
C	34	245.6	6.2	969	13	EX415714	EX415714
C	35	245.6	6.2	2284	11	BC036238	Homo sapi
	36	245.4	6.2	467	28	AQ388230	RPC111-14
C	37	245.4	6.2	785	12	BM721013	UI-E-EO0-
	38	244.8	6.2	663	29	AG174449	AG174449
C	39	244.6	6.2	750	12	BM678642	UI-E-BO0-
C	40	244.2	6.2	1201	13	EX385602	EX385602
	41	244	6.2	548	14	CB144391	K-EST0198
	42	243.8	6.2	755	28	BZ599689	WHABO08TR
C	43	243.8	6.2	2971	28	AF101960	AF101960
	44	243	6.1	826	28	BZ611620	WHADA597F
C	45	243	6.1	929	9	AL580585	AL580585

ALIGNMENTS

RESULT 1
BM563325
LOCUS
DEFINITION
AGENCOURT_6564864 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5743936
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BM563325
AGENCOURT_6564864 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5743936
5', mRNA sequence.
BM563325
BM563325.1 GI:18810153
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1238)
NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: LLAM12764 row: h column: 17
High quality sequence stop: 622.

FEATURES	Location/Qualifiers
source	1..1238
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:5743936"
	/tissue_type="medulla"
	/lab_host="DH10B"
	/clone_lib="NIH_MGC_119"
	/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."
ORIGIN	
Query Match	16.4%; Score 649.4; DB 12; Length 1238;
Best Local Similarity	98.7%; Pred. No. 2.9e-90;
Matches 686; Conservative	0; Mismatches 6; Indels 3; Gaps 3;
QY	3266 TGGGGTGGCTGCCACGCTCCCTGCCACGCGCTTGGGGGACGGGAGACGGGACGGAGA 3325
Db	20 TGGGGTGGCTGCCACGCTCCCTGCC-CGGGCTGGGGGACGGGAGACGGGACGGAGA 78
QY	3326 TGTTAGTGTGGGGCCGCCCGAGGGTTCAACACTGTTTCTGAGAAACTTCCCACTGC 3385
Db	79 TGTTAGTGTGGGGCCGCCCGAGGGTTCAACACTGTTTCTGAGAAACTTCCCACTGC 138
QY	3386 CCACCCACCCGTTCTCCGTGTGCCGAGGCCGGTCTCGGGTAGGCTCCGCCGCCGAGC 3445
Db	139 CCACCCACCCGTTCTCCGTGTGCCGAGGCCGGTCTCGGGTAGGCTCCGCCGCCGAGC 198
QY	3446 CCCAAACCGGGTCCCGACGCCCTTCCAGAGAGAAAGCTCCCGACCGGGATGCGGGCAG 3505
Db	199 CCCAAACCGGGTCCCGACGCCCTTCCAGAGAGAAAGCTCCCGACCGGGATGCGGGCAG 258
QY	3506 AGGCCACAGCGCGGTGGAAGAGAGAGCTGAGAGAGGAGAAACAGAGGGAGGGGGAGCGAG 3565
Db	259 AGGCCACAGCGCGGTGGAAGAGAGCTGAGAGAGGAGAAACAGAGGGGAGGGGGAGCGAG 318
QY	3566 GAGCTGGCGGCAGAGGGAAACAGCAGATTGGCCGAGCGCAATGCAACGGCAGACGACGAT 3625
Db	319 GAGCTGGCGGCAGAGGGAAACAGCAGATTGGCCGAGCGCAATGCAACGGCAGACGACGAT 378
QY	3626 GGCACCAATTCCTTTCCGCCAATGACGAGCCGGAGTTTACAGAGCCTCATTTAGCATTT 3685
Db	379 GGCACCAATTCCTTTCCGCCAATGACGAGCCGGAGTTTACAGAGCCTCATTTAGCATTT 438
QY	3686 CCCACAGACGAGGGGACGAGGCGAGGCGGGTGTGTGGTGTGCGTGTGGCAGCATCC 3745
Db	439 CCCACAGGACGAGGGGACGAGGCGAGGCGGGTGTGTGGTGTGCGTGTGGCAGCATCC 498
QY	3746 CCGGCGCCCTTGCTGCGGTCCGCCGAGCCTTCGGCCTCTGTCTCCT-CCCCCTCCCGCCCT 3804
Db	499 CCGGCGCCCTTGCTGCGGTCCGCCGAGCCTTCGGCCTCTGTCTCCTCCCGCCTCCCGCCT 558
QY	3805 TACCTTCAGCGGGACCGCCCGGCGCAGTCAACTCTTCGCACTTTGCGCCCTGTTGGCAG 3864
Db	559 TACCTTCAGCGGGACCGCCCGGCGCAGTCAACTCTTCGCACTTTGCGCCCTGTTGGCAG 618
QY	3865 CGGATAAAGGGGGCTGAGGAATAACCGGACCGGTCAACCGGTGCGAGCTCTAGCCTTT 3924
Db	619 CGGATAAAGGGGGCTGAGGAATAACCGGACCGGTCAACCGGTGCGAGCTCTAGCCTTT 678
QY	3925 AAATTCCC-GGCTCGGGGACCTTCCACGACCGCGG 3958
b	679 TAATTCCCGGGCTTCGGGGACCTTTCGGGACCGGGG 713

```

RESULT 2
BI559696
LOCUS
DEFINITION
603252566F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5294921 5',
mRNA sequence.
ACCESSION
BI559696
VERSION
BI559696.1 GI:15447010
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 579)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1A11746 row: C column: 18
High quality sequence stop: 576.
Location/Qualifiers
1..579
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5294921"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
ctcgag; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to 200 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH MGC Library."
FEATURES
source

```

ORIGIN	Query Match	12.4%;	Score 491.4;	DB 12;	Length 579;
	Best Local Similarity	97.6%;	Pred. No. 8.4e-66;		
	Matches 562;	Conservative 0;	Mismatches 7;	Indels 7;	Gaps 6;
QY	3229	GGAGGGGGGGCTTCCGGCTCCCGCTCCCAAGCTCCAGATCCTGGGGTGGTGCCACGTCCTCCCT	3288		
Db	4	GGAGTCGGGGCTTCCGGCTCCCGCTCCCAAGCTCCAGATCCTGGGGTGGTGCCACGTCCTCCCT	63		
QY	3289	GCACGCGGCTGGGGGAGCGGAAGACGGAGCGGAGATGTTAGT-GGTGGGGCCCCCG	3347		
Db	64	GCACGCGGCTGGGGGAGCGGAAGACGGAGCGGAGATGTTAGTGGGTGGGGCCCCCG	123		
QY	3348	AGGG--TTCACCACTGTTTCTTGAGAAACTTCCCAAGTCGCCACACCCAGTCCTCCGTG	3405		
Db	124	AGGGGTTCAACNCACTGTTTCTTGAGAAACTTCCCAAGTCGCCACACCCAGTCCTCCGTG	183		
QY	3406	TGCCGAGGCGCGTCTGGGTAGGTCCTGGCCCCAGCCCCAAACCGGGTCCCCAGCC	3465		
Db	184	TGCCGAGGCGCGTCTGGGTAGGTCCTGGCCCCAGCCCCAAACCGGGTCCCCAGCC	243		
QY	3466	CTTCTCAGAGGAAAGCTCCCGACCGGGATCGGGGACAGGCCACGCGGGGTGGA	3525		
Db	244	CTTCTCAGAGGAAAGCTCCCGACCGGGATCGGGGACAGGCCACGCGGGGTGGA	303		
QY	3526	GAGAAGCTCAGAGGAGAAACAGAGGGGAGGG-GGAGCGAGGAGCTGGCGGACAGGGAA	3584		

full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

```

ORIGIN
Query Match      8.4%; Score 334; DB 12; Length 754;
Best Local Similarity 99.7%; Pred. No. 1.1e-41;
Matches 345; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 3613 GGCAGGACGAGTGGCCAAATTCCTTCGGCCCAATGACGAGCGGAGTTTACAGAAGC 3672
Dbb|||||
Dbb1 GGCAGGACGAGTGGCCAAATTCCTTCGGCCCAATGACGAGCGGAGTTTACAGAAGC 60
QY 3673 CTCATTAGCATTTCCAGAGCGAGGCGGAGGCGGAGGCGGAGGCGGAGTGTGTGCGGT 3732
Dbb|||||
Dbb61 CTCATTAGCATTTCCAGAGCGGAGGCGGAGGCGGAGGCGGAGTGTGTGCGGT 119
QY 3733 GTCGCGAGCATCCCGCGCGCCCTGCTCGGFCGCGGAGGCTCGGCTCTGTCTCTCC 3792
Dbb|||||
Dbb120 GTCGCGAGCATCCCGCGCGCCCTGCTCGGFCGCGGAGGCTCGGCTCTGTCTCTCC 179
QY 3793 CCCTCCCGCCTTACCTTCACGCGGAGCGCGCGCGGAGTCAACTCTCTGCACTTTGCC 3852
Dbb|||||
Dbb180 CCCTCCCGCCTTACCTTCACGCGGAGCGCGCGCGGAGTCAACTCTCTGCACTTTGCC 239
QY 3853 CTGCTTGGCAGCGGATTAAGGGGGCTGAGGAATACCGACACGCTACCGTTGCCA 3912
Dbb|||||
Dbb240 CTGCTTGGCAGCGGATTAAGGGGGCTGAGGAATACCGACACGCTACCGTTGCCA 299
QY 3913 GCTCTAGCTTTAAATTCGCGGCTCGGGGACCTCCACGACCGCGG 3958
Dbb|||||
Dbb300 GCTCTAGCTTTAAATTCGCGGCTCGGGGACCTCCACGACCGCGG 345

```

```

RESULT 5
BC035771/c
LOCUS
DEFINITION
Homo sapiens RAD1 homolog (S. pombe), mRNA (cDNA clone
IMAGE:5582524), with apparent retained intron.
ACCESSION
BC035771
VERSION
BC035771.1 GI:23958889
KEYWORDS
HTC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1910)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shvachenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,U.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
REFERENCE
1 (bases 1 to 1910)
Strausberg,R.
Direct Submission
TITLE

```

JOURNAL

REMARK
COMMENT

Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-i@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nhgri.nih.gov
 Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Teurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 79 Row: d Column: 15
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 this clone has the following problem: retained intron.

FEATURES

Location/Qualifiers

1..1910

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5582524"
/tissue_type="Testis, embryonal carcinoma"
/clone_lib="NIH MGC_92"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

```

ORIGIN

```

Query Match      7.4%; Score 294.2; DB 11; Length 1910;
Best Local Similarity 64.5%; Pred. No. 8.2e-36;
Matches 496; Conservative 0; Mismatches 253; Indels 20; Gaps 3;

QY 1414 TGTAAACAAATAATAAATAATCAAAATTTTTTTTTTTTGTGAGAAAGAGTCTGTCTC 1473
Dbb|||||
Dbb1176 TGCAGCTAAACAAATAATTTTTTTTATTTTATTTTGTGAGCGAGTCTGTCTC 1117
QY 1474 TTTACCCCTGGCTGGAGGCGACTGGTGTGATCTCAGCTCAGTCAGCACTTCGCCCTCCCGG 1533
Dbb|||||
Dbb1116 TGTACCCAGGCTGGAGTGCATGTGCATCTCGGCTCAGTCAACCCCGCCCTTCAG 1057
QY 1534 GTTCAAGCAATTCCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAGCACTTCCACC 1593
Dbb|||||
Dbb1056 TTTCAGTGATTTCTCTGCTCAGCCTCTCAAGTAGCTGGGATTACAGACACCCACC 997
QY 1594 ATGCCCGAGCTGATTTTCTATTTTGTAGTAGAT--GGGATTTCACTTTTGGCCCAAGCT 1652
Dbb|||||
Dbb996 ACACCCAGCTAATTTTGTATTTTGTAGAGGTAGGGGTTTTCACCATGTTGGCCAGGCT 937
QY 1653 GGCTCTCAAACTTTTGTCTGTCTAATTTTGTAACTATTGTTCTTTTGTGCTGAGTAGGG 1712
Dbb|||||
Dbb936 GATCTCGAATCTCTGACCTCAGA-----TGATCCCTCGCTCGGC 895
QY 1713 CCCCGACACCAAAATAAATAATCTTAGAATCCAAATCAGTGTGTGTGTGACCACTGT 1772
Dbb|||||
Dbb894 TTCCCAAAATGCTGGGATTACAGGTGTGAGCCACCACTCCCGGCCCTAAATAATAGTTT 835
QY 1773 CACTGTGACACCAAGTGTGACCGGGCTCAGGAGTAGAGTGTCTCTGCTCGAAGA 1832
Dbb|||||

```

Db 834 AAATGTACAAAATAATCCATCCTTACTGAGCTGATTTATATTTGTTTAAATCTTTTAA 775
 QY 1833 GAAATAGAAATG-AAAATATTTCTCGGGCCAGCGGTGGTGGTCTCATCGCTGTAATCCAGC 1891
 Db 774 GATTCAACTTTTAAATAATATTAGCCAGGATGATAGCTCACGCCCTGTAATCCAGC 715
 QY 1892 ACTTTGGAGGCCAAGGATGTGATCAGCTGAGGTCAGGAGTTCAAAACCCAGCTGGCC 1951
 Db 714 CTTTTGGAGGCCAGCGCGGTGGATTGCTTGAGCTCAGGAGTTCAAGATTAGCTAGGG 655
 QY 1952 AACATGCTGAACCCCGTCTCTACTAAATAACAAAATTAAGTAACTAGCTAGTGGTGGCA 2011
 Db 654 AACATAGTAAACCCCATCTCTACAAAATAATTAAGGAGTAACTAGGAGTGGAGCATG 595
 QY 2012 TGCTGTAAATCCAGCTACTTGGAGGGTGAGCGAGAGAAATTTCTTGAACCCGGAGGC 2071
 Db 594 CACCTGTAGCTAGCTATTTGGAGCTGAGATGGAGGATGCTTGAGCTTAGAGAT 535
 QY 2072 AGAGTTGAGTGAAGCGAGATCACACCTGCTGCTGAGCTGCTGAGGAGAGAGAGAC 2131
 Db 534 TGAGGCTGAGTGAAGTGAATCTTACAACTGCTGCTGAGCTGCTGAGGAGAGAGAT 475
 QY 2132 TTCTCTCAAAAACAAAACAAAAGATTAAGCAATTTAGACATTG 2180
 Db 474 CTGTCTCCAAAATAAAATAAAACTGAGATCAAAATAAATTAG 426

RESULT 6
 LOCUS BC035510
 DEFINITION Homo sapiens sialophorin (gpl15, leukostalin, CD43), mRNA (cdna clone IMAGE:5094058), with apparent retained intron.
 ACCESSION BC035510
 VERSION BC035510.1 GI:23271054
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2076)
 Straubeberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Schmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, X.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 2238257
 MEDLINE 12477932
 PUBLISHED 2 (bases 1 to 2076)
 Straubeberg, R.
 Direct Submission
 Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lari, P., Legaspi, R., McDowell, J., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., Mudge, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsourgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 39 Row: p Column: 13
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
 1..2076
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5094058"
 /tissue_type="Lymph, Burkitt lymphoma"
 /clone_lib="NIH MGC 8"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"

ORIGIN

Query Match 7.3%; Score 250.6; DB 11; Length 2076;
 Best Local Similarity 67.2%; Pred. No. 2.8e-35;
 Matches 498; Conservative 0; Mismatches 214; Indels 29; Gaps 5;
 QY 1421 AAAATAATAAATAATACAAATTTTTTTTTTTTGTGAGAAAGAGCTTTGCTTTTACC 1480
 Db 741 ACAACGTTGAATAGATGCTTTTTTTTTTTTGTGAGATGGAGTTTCACTCTTGTACC 800
 QY 1481 CTGGCTGAGGCGAGTGGTGTGATCTAGCTCAGTCACTGAACTTCCCTCCCGGTTCAAG 1540
 Db 801 CAGGCTGAGTGAATGGCTGATCTTGGCTCACCAGAACCTCCGCTCTCTGGTTCAA 860
 QY 1541 CAATTCTCTCTTCCAGCTCCCAAGTAGCTGGGACTACAGCACTCCCAACCATGCCCA 1600
 Db 861 CAATTCTCTCTTCCAGCTCCCAAGTAGCTGGGACTACAGCACTCCCAACCATGCCCA 920
 QY 1601 GCTGATTTTGTATTTTGTAGTGGAGTGGGATTTTCACTTTTGTGGCCAAAGCTGTTCAA 1660
 Db 921 GAT-AAATTTTGTATTTTGTAGTGGAGTGGGATTTTCACTTTGTGGCTGAGTCTTGA 979
 QY 1661 ACTTTTGTCTCATATTTGTGTAACTATTTCTTTTCTTTTGTGAGTGGGCCCCAGA 1720
 Db 980 ACT-----CCGAGCTCTGCTGATCCGCTATGCTCGGCTCCCAAGTCTGGGATT 1031
 QY 1721 CCATAAATAATAATCTTAGAATCAATCACTGTGTGTGTGACCACTGTCACTTGA 1780
 Db 1032 ACATGCTGTAGCCACCCTGCGCCGACATAGATGCTTTTAAATTTTCTGGAGAAAGCA 1091
 QY 1781 AACACAGTGTGACCGAGGCTCAGGAGTAGAGTGATCTCTGCTCGAAAGAGAAATAGA 1840
 Db 1092 AAGCAAAAGAACGAGTGGATATTTTAA-----ACTAAAGAGAAACAAA 1137
 QY 1841 ATGAAATATTTCTCCGGCCAGGCGTGGTCTC-ATGCCGTGTAATCCACACTTTGGG 1899
 Db 1138 AAGAGGAGTAGACGAGCGCAGAGTGGTGGCTCAACGCTGTGTATCCACACTTTGGG 1197
 QY 1900 AGGCCAAGGATGTGGATCACTGAGTGCAGGAGTTCAAAACCAACGCTGGCCACATGGT 1959


```

Db      1198 AGCCGAGGCGAGTGGATCACCTGAGTCAAGAGCCAGGCTCAACCAATGGT 1257
QY      1960 GAAACCCGCTCTCTACTAAAAATACAAAAATTTAGCTAGTGTGGGCGCATGCTGTGTA 2019
Db      1258 GAAACCCGCTGTTC-----AAATACAAAAATTTAGCTGGCGGTGGGCGGCACCTGTG 1312
QY      2020 ATCCAGCTACTTTGGAGGCTGAGGCGAGGAGAAATTTCTGAACCCCGGAGGCGAGAGTTG 2079
Db      1313 ATCCAGCTACTTTGGAGGCTGAGGCGAGGAGAAATCCCTTGAACCCAGGAGGTGGAGTTG 1372
QY      2080 CAGTGAAGCGAGATCACACACTGCTCCTCAGCTGGGCGAGAGCGAGATTCCTCTC 2139
Db      1373 CAGTGAAGCGAGATCACGCAATTCGCACTCAGCTGGGCGAGCAAGTAAAAAACTCCATCT 1432
QY      2140 AAAAAACAAAAACAAAAAGA 2160
Db      1433 CAAAAAAGAGGATA 1453

BC038630      2772 bp mRNA linear HTC 04-MAR-2003
Homo sapiens, Similar to hypothetical protein FLJ20489, clone
IMAGE:5263792, mRNA.
ACCESSION      BC038630
VERSION        BC038630.1 GI:24116283
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
REFERENCE
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE
JOURNAL      Strausberg, R.
              Direct Submission
              Submitted (15-OCT-2002) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK
COMMENT      NIH-MGC Project URL: http://mgc.nci.nih.gov
              Contact: MGC help desk
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
              cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
              Toshiyuki and Piero Carninci (RIKEN)
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Institute for Systems Biology
              http://www.systemsbio.org
              Contact: amadan@systemsbiology.org
              Anup Madan, Jessica Fahey, Erin Helton, Mark Ketterman, Anuradha
              Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 73 Row: i Column: 4
This clone has the following problem: retained intron.
Location/Qualifiers
1..2772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5263792"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"

FEATURES
source
ORIGIN
Query Match 7.3%; Score 290.6; DB 11; Length 2772;
Best Local Similarity 69.5%; Pred. No. 2.3e-35;
Matches 508; Conservative 0; Mismatches 189; Indels 34; Gaps 7;
QY 1441 TTTTITTTTTTTTTTGAGAAAGCTCTTGCTCTT-TCACCTGGCTGGAGGCGAGTGGT 1499

```

```

Db      2753 TTTTITTTTTTTTTTGTAGACGAGTTTGTCTCTTGTCTCCAGGCTGGAGTGCATGGT 2694
QY      1500 GTGATCTCAGCTCAGTGCNAACCTTTCGCTCCCGGGTTCAAGCAATTCCTCTGTTGAGCC 1559
Db      2693 GCCATCTCGGCTCGCTGCAAACTGCACCTCCCGGGTTCAACAGATTCCTCTGCTCAGCC 2634
QY      1560 TCCCAAGTAGCTGGGACTACAGGCACTTCCCAACCTGCTCCAGCTGATTTTGTATTATTTTA 1619
Db      2633 TCCCAAGTAGCCCGGACCAAGTGGCGCCACAGCCGCCCAATTTCTGTATATTTTA 2574
QY      1620 TTAGAGATGGGATTTTCACTTTGTGGCCCAAGTGGTCTCAAACTTTTGTCTGCTCAATTT 1679
Db      2573 GTATATTTGGGGTTTTCACCATGTTGSCCAGGCTGCTCAAACTCCTGACCTCAGGTGAT 2514
QY      1680 GTTGTAACTATTTGTTCTTTTGTCTGAGGTAGGCGCCCGACCAAAAAAATAATCTTTA 1739
Db      2513 CCACCTCGCTCAGCCTCCCAAGGTGGTG-----GGATTACAGGCATGAGCCACA 2465
QY      1740 GAATCCAAATCAGTGTGTTGGTTTGACCACTCTCACTTGAGAACCACTGTGACCGGG 1799
Db      2464 GCTCCAGCTAGGATTTCTTGATATATCAACTATCTTGTGTAG-AGTGAATCTGGC 2406
QY      1800 CTTAGGAGTAGAGGTGATCTCTGCTCGAAAGAGAAATAGAAATATTTCTCCGGGC 1859
Db      2405 TAAATTAATAAGAGATGAATTTACTGAAAGTGTATA-----TGC 2364
QY      1860 CAGGCTGTGGTCTCATGCTGTATATCCAGCACTTTGGAGGCCAAGGCATGTGATCA 1919
Db      2363 CAGGAGTGTGGCTCATGCTGTATATCAACACCTTGAGAGCGCGGTTAGGTCA 2304
QY      1920 CTTGAGGTGAGGAGTTCAAAACCCAGCTGGCCAACTAGTGAACCCCGTCTCTACTAAA 1979
Db      2303 CTTGAGGTGAGGAGTTC-AAACCACTGACCAACTGTTGAAACTTGCCTCTACTAAA 2245
QY      1980 AATCAAAAAATTAGCTAAGTGTGGCGCATGCTGTAAATCCAGCTACTTTGGAGGG 2039
Db      2244 AATAC-AAAAATCAGCCAGGCGTAGTGTGATGCTGTAGACCCAGCTACTCTCCGAGGC 2186
QY      2040 TGAGCAGGAGAAATTTCTTGAACCCGCGGAGGAGAGTTGCAGTGAAGCGAGATCACACC 2099
Db      2185 TGAGCAGGAGAAATTTGCTTGAACCCAGGAGGTAGAGTTGCAGTGAAGCGAGATCACACC 2126
QY      2100 ACTGCACCTCCAGCTTGGGCGGAGAGAGCACTTCTCTCAAAAAAACAACAAAAAAG 2159
Db      2125 ATTGACT-CAGCCTGGCAACAGAGGAGAGTCCATCTCAAAAAAAGAAAGAAAGRA 2067
QY      2160 AATTAAGCAAA 2170
Db      2066 AAGAAAAGAAA 2056

BC038630      2772 bp mRNA linear HTC 04-MAR-2003
Homo sapiens, Similar to hypothetical protein FLJ20489, clone
IMAGE:5263792, mRNA.
ACCESSION      BC038630
VERSION        BC038630.1 GI:24116283
KEYWORDS      HTC.
SOURCE
ORGANISM      Homo sapiens (human)
REFERENCE
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE
JOURNAL      Strausberg, R.
              Direct Submission
              Submitted (15-OCT-2002) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK
COMMENT      NIH-MGC Project URL: http://mgc.nci.nih.gov
              Contact: MGC help desk

```

Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 Contact: amadan@systemsbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 73 Row: i Column: 4
 This clone has the following problem: retained intron.

FEATURES

source
 1. 2772
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5263792"
 /tissue type="Brain, hippocampus"
 /clone_lib="NIH MGC_95"
 /lab host="DH10B"
 /note="Vector: pBluescript"

ORIGIN

Query Match 7.2%; Score 283.4; DB 11; Length 2772;
 Best Local Similarity 66.1%; Pred. No. 3e-34;
 Matches 486; Conservative 0; Mismatches 216; Indels 33; Gaps 4;
 QY 1441 TTTTCTTTTCTTTTGTGAGAAAGTCTTGTCTTTCACCCCTGGCTGGAGGCGAGTGTG 1500
 DB 2071 TTTCTTTTCTTTTCTTGTGAGATGGAGTCTTCTCTGTGTCGCCAGCT-GAGTGCATGGTG 2129
 QY 1501 TGATCTCAGTCATCGAACTTTCGCTCCCGGTTCAAGCAATTCCTGCTTCAGCCT 1560
 DB 2130 TGATCTCGGCTCATCGAACTTTCACCTCTCTGCTGGTTCAGCAATTCCTGCTTCAGCCT 2189
 QY 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCCACCATGCCAGCTGATTTTGTATTTTAG 1620
 DB 2190 CCGGAGTAGCTGGTCTACAGGATGACACACTAGCCTGGCTGATTTTGTATTTTAG 2249
 QY 1621 TAGAGTAGGATTCATTTTGTGGCCAACTGGTCTTCAAACTTTTGTCTCATATG 1680
 DB 2250 TAGAGGCAAGTTTCCACATGTTGGTCAGGCTGGTTGAACCTCTGACCTCAGTGCACCT 2309
 QY 1681 TTGTAACTATTTCTTTTCTGAGTAGGGCCCCCAGACCAACCAATAAATCTTAG 1740
 DB 2310 ACCGGCTCGGCTCTCAAGGTGTGTGATTAAGGATGAGCCACCACTCTCGGCATAT 2369
 QY 1741 AATCCAAATCAGTGTGTGTTTGTGACCACTGTCTAGAACCAACAGAGTGACAGGGC 1800
 DB 2370 AACACTTTTCAGTAAAT-----CATCTCTATTATTAGCCAGAGTTCACCTCTA 2422
 QY 1801 CTCAGAGTAGAGGTGATCTCTCTCGAAGAGAAATAGAAATATTCTTCGGGCC 1860
 DB 2423 ACAAGCAAGATAGTTGATATATCAAGAA-----TCCTAGCT 2458
 QY 1861 AGCGTGTGGCTCATGCTGTAATCCAGCACTTTGGGAGGCCAAGGCAATGGAATCAG 1920
 DB 2459 GAGAGCTGGCTCATGCTGTAATCCACCACTTGGGAGGCTGAGCGAGTGATCAG 2518
 QY 1921 CTGAGTCCAGGAGTTCAAAACCAAGCTGGCCAACTGGTGAACACCCGCTCTTACTAAAA 1980
 DB 2519 CTGAGTCCAGGAGTTTGTAGACCAAGCTGGCCCAATGTTGAACCCCAATATCTAAAT 2578
 QY 1981 ATACAAAAATTAGCTAAGTGTGGTGGCGCATGCTGTAAATCCAGCTACTTGGGAGGT 2040
 DB 2579 ATAC-AGAAATTGGCTGGGCGTGGTGGCGGCCCACTGTGCTCCGCTACTTGGGAGGT 2637
 QY 2041 GAGGCGAGGAGATTTCTTCAACCCGGAGGCGAGAGGTTCCAGTGAAGCAGATCACCA 2100
 |||||

Db 2638 GAGGCAGGAGATCTGTTCAACCCGGAGGTGAGTTCGAGCGAGCCGAGATGGCACCA 2697
 QY 2101 CTGACTCCAGCTCGGGGAGAGAGCGAGACTTCCTCTCAAAAAACAAAAACAAAGA 2160
 Db 2698 TTGCACTCCAGCTCGGAGACAGAGCAAACTCGTCTTACCAAAAAAAGAAAGA 2757
 QY 2161 ATTAGCAAAATTAGA 2175
 Db 2758 AAAAAAAGAAAAA 2772

RESULT 9

BC035510/c

LOCUS BC035510

DEFINITION Homo sapiens sialophorin (gpl15, leukosialin, CD43), mRNA (cdna

clone IMAGE:5094058), with apparent retained intron.

ACCESSION BC035510

VERSION BC035510.1

KEYWORDS GI:23271054

SOURCE HTC

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 2076)

AUTHORS Mammalia; Eutheria; Primates; Carnivora; Vertebrata; Euteleostomi;
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,
 Aramson, K.J., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smallos, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nisc.nih.gov

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-I., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 39 Row: p Column: 13
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 This clone has the following problem: retained intron.

FEATURES
 source
 1..2076
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5094058"
 /tissue_type="Lymph. Burkitt lymphoma"
 /clone_lib="NIH_MGC_8"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"

ORIGIN

Query Match 6.9%; Score 273.4; DB 11; Length 2076;
 Best Local Similarity 67.4%; Pred. No. 1.2e-32;
 Matches 487; Conservative 0; Mismatches 206; Indels 30; Gaps 6;
 QY 1444 TTTTCTTTTTCAGAAAGHCTTCTCT-TCACCTGCTGGAGGCGAGTGTGTG 1502
 Db 1446 TTTTCTTTTTCAGATGGAGTTTTCCTTTGTCGCCAGGCTGGAGTGCATGGCGTG 1387
 QY 1503 ATCTCAGCTCACTGCACTTTCGCTCCGGGTTCAAGCAATCTCTGCTTCAGCTCC 1562
 Db 1386 ATCTCGGCTCACTGCAACCTCCACTCTCGGTTCAAGGATTCCTGCTCAGCTCC 1327
 QY 1563 CAAGTAGCTGGGACTACAGGCACTTCCACATGCGCCAGCTGATTTTGTATTTAGTA 1622
 Db 1326 CAAGTAGCTGGATCACAGTGCCTGCCACACAGCCAGCTAATTTTGTATTTT 1271
 QY 1623 GAGATGGATTTCACTTTGTGGCCAGCTGGTTCATAAATTTTGTGTCATAAATGTT 1682
 Db 1270 GAAACAGGGTTTCCACATGTTGTTCAGCTGGTCTTGAATCT-----CCTGACCTCAG 1219
 QY 1683 GTAACATATTTGCTTTGCTAGGTAGGCGCCGACCAACAAATAAATAAATCTTAGAA 1742
 Db 1218 GTGATCCACTGCTCGGCTCCCAAGTGTCTGGGATTACAGAGCTTGAGCCACCACTG 1159
 QY 1743 TCCAAATCAGTGTGTGTTTGACCACTGTCAGTGCAGAACCACTGTCAGGAGGCT 1802
 Db 1158 TGCCCTGCTATCTCTCTTTTGTGTTTCTTTTGTGTTTGTGTTTGTGTTTGTCT 1099
 QY 1803 CAGGAGTAGAGTGATCTCTGCTCGAAAGAGAAATAGAAATATTTCTCGGGCCAG 1862
 Db 1098 TTTGCTTTGCTTTTCTCTCC-----AGAAATTAAGACATCTATTGTCGG 1052
 QY 1863 GCGTGTGCTCATGCTCTAATCCAGCACTTTGGGAGCCAGGCAATGTCATCACT 1922
 Db 1051 GCACGTGCTCAGCATGATATCCAGCACTTTGGGAGCGGAGGATACAGGATCAGCA 992
 QY 1923 GAGGTGAGGATTCAAAACAGCTGGCCAAATGTTGAAACCCGCTCTCTACTAAAAAT 1982
 Db 991 GAGGTGCGGAGTTCAGACAGGCTGACCAATGTTGAAACCCCATCTCTACTAAAAAT 932
 QY 1983 ACAAAAATATGCTAAGTGTGGCGCATGCTGTAAATCCAGCTACTTGGGAGGTTGA 2042
 Db 931 AC--AAAAATATCCGGGTTTGTGTGTCATGCTGTATCCAGCTACCTGGGAGGCTGA 874
 QY 2043 GCGAGGAGATTTCTTGACCGGAGGCGAGGTTGAGTGAAGGAGATCACACCTACT 2102
 Db 873 GCGAGGAGATTTGTTGAACCGGAGGCGGAGGTTCTGGTGAAGCAAGATCACGCCATT 814
 QY 2103 GCATCCAGCTGGGGAGA--GAGCGAGACTTCTCTCAAAAAACAAAAACAAAGA 2160
 Db 813 GCATCCAGCTGGGTTAACAAGTGAATCTCTCAATCTCAAAAAACAAAAACAAAGCATCT 754
 QY 2161 ATT 2163

Db 753 ATT 751

RESULT 10
 BUI75345

LOCUS
 DEFINITION
 BUI75345
 5', mRNA sequence.
 BUI75345
 EST.
 BUI75345.1. GI:22689329

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-rc@mail.nih.gov
 Tissue Procurement: ATCC

FEATURES
 source

1..749
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:601997"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_92"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
 Average insert size 2.5 Kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 6.8%; Score 270; DB 13; Length 749;
 Best Local Similarity 64.5%; Pred. No. 7.4e-32;
 Matches 462; Conservative 0; Mismatches 230; Indels 24; Gaps 3;
 QY 1443 TTTTCTTTTTCAGAAAGTCTTCTCTTTTCCCTGCTGGAGGCGAGTGTGTG 1502
 Db 52 TTTTCTTTTTCAGAAAGTCTCTACTCTGTCAACCCAGGCTGGAGTTCAGTGTGCC 111
 QY 1503 ATCTCAGCTCACTGCACTTTCGCTCCGGGTTCAAGCAATCTCTGCTTCAGCTCC 1562
 Db 112 ATCTTGGCTCACTGCAACTCTGCTCTCTGCTTCCAGTATCTCCCATCTTAGCTCC 171
 QY 1563 CAAGTAGCTGGGACTACAGGCACTTCCACCATGCCAGCTGATTTTGTATTTAGTA 1622
 Db 172 TGAGTAGCTGGGACACAGGCGTGTGCCACACATCCAGCTAATTTTGTATTTTGTG 231
 QY 1623 GAGATGGATTTCACTTTTGTGGCAAGCTGTCTCAAACTTTTGTCTGTCTATAATTGT 1682
 Db 232 GAGACGGGTTTTCATCATGTTTGGCAGGCTGTGTGCACTCTCTGAC----- 278
 QY 1683 GTAACATTTGTTCTTTTGTGAGTGGGCCCCCAGACCAAAAAATAAATCTTAGAA 1742
 Db 279 -----CTAAGAGCATCTGTCTGCTCGGCTCCCAAGTGTGGGTAAGTCTGGGAT 332
 QY 1743 TCCAAATCAGTGTGTGTTTGGACCATGTCTACATTTGAGAACCACTGTGACCGGCT 1802

	Query Match	6.8%	Score 269.4;	DB 12;	Length 1172;
	Best Local Similarity	97.8%;	Pred. No. 7.1e-32;		
	Matches 273; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0;
QY	3680	GCATTTTCCCGAGGCGACGGCGCAGAGCCCGGTGTGTGTGTCGGTGTCGGCA	3739		
Db	6	GGAAATTCTCGGNATCAGGGGCAGGGCAGAGCCCGGTGTGTGTGTCGGTGTCGGCA	65		
QY	3740	GCATCCCGCGGCCCTGTCTCGGTTCGGTCGGCGGAGCCTCGGCTCTGTCTCTCCGCCCTCCC	3799		
Db	66	GCATCCCGCGGCCCTGTCTCGGTTCGGTCGGCGAGCCTCGGCTCTGTCTCTCCGCCCTCCC	125		
QY	3800	GCCTTTTACTTCCAACGCGGACCCCGCGCGCAGTCAAACCTCCTCGCATTTTGGCCCCCTGTT	3859		
Db	126	GCCTTTTACTTCCAACGCGGACCCCGCGCGCAGTCAAACCTCCTCGCATTTTGGCCCCCTGTT	185		
QY	3860	GCACGGGATAAAGGGGCTGAGGAATACCGGACACGGTCA CGGTTCGGCAGCTTAG	3919		
Db	186	GGCAGCGGATAAAGGGGG-TGAGGAATAACCGGACACGGTCA CCGGTTCGGCAGCTTAG	245		
QY	3920	CCTTTAAATTCGCGGCTCGGGACCTCCACGCCACCGGG	3958		
Db	246	CCTTTAAATTCGCGGCTCGGGACCTCCACGCCACCGGG	284		

RESULT 12	BC035771	LOCUS	BC035771	1910 bp	mRNA	linear	HTC 19-NOV-2003
DEFINITION	Homo sapiens RAD1 homolog (S. pombe), mRNA (cDNA clone IMAGE:582524), with apparent retained intron.						

BC035771.1 GI:23536883

VERSIONS	ORGANISM	REFERENCE
HTC.	Homo sapiens (human)	
GENOTYPE	Homo sapiens	
SOURCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	1 (bases 1 to 1910)	
	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altchul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Udwin,T.B., Toshiyuki,S., Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaby,S.J., Bosak,S.A., McSwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worthy,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kerteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.	

TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE	22398257
PUBMED	12477932
REFERENCE	2. (bases 1 to 1910)
AUTHORS	Strausberg, R.
JOURNAL	Direct Submission
TITLE	Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgapbs@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc.

Email: cgapds-1@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.

Query Match	6.6%	Score 259.4	DB 9	Length 929
Best Local Similarity	62.2%	Pred. No. 2.8e-30		
Matches 507	Conservative 14	Mismatches 245	Indels 49	Gaps 6

QY	1441	TTTTTTTTTTTTTTTGAGGAAGAGCTCTGCTCTTTCACCTCGCTGGAGGGCAGTGCTG	1500
Db	116	TTTTTTTTTTTTTTTGTAGTCGAGATCTCTCCCTGTGCGCAGGCTGAGACAGTGGCG	175
QY	1501	TGATCTCAGCTCATGTCAACTTTCGCCTCCCGGGTTCAAGCAAAATCTCCTGCTTCAAGCCT	1560

```
Db 176 TGATVTCAGCTCGCTCAACCTCTCCCTCCCGTGTTCAGCAATTTCTCTGCTCAGCCT 235
QY 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCACCATGCCAGCTGATTTTGTATTTTATAG 1620
Db 236 CCCGAGTGWCTGGGACTACAGGCGTGTCCACCACTCAGCTAATTTTGTATTTCCAG 295
QY 1621 TAGCATCGGATTTTCACTTTGTGGCCAGCTGGTCTCAAACTTTTGTGTGCATATTTG 1680
Db 296 TAGAGACTG--TTTCACCATGTTTGTGTGAGCTGGTCTTGAACCTCTGACTTCGTGA--- 349
QY 1681 TTGTAACTATTGTTCTTTTGTGAGGTAGGGCCCCAGACCAAAAAATAAATCTTAG 1740
Db 350 -----TCCGCCACCTCAGCCTCCCAAGTGTGGGATTACAGCGCAG 392
QY 1741 AATCAAAATCAGTGTGGTGTGAGTACCTGTCACATGAGAACACAGTGTGACCGAGGC 1800
Db 393 CCACC-----TTGCCCGCGCGAAGTTAAAGATGCGAGTTGGGATAAGACC 437
QY 1801 CTCGAGGTAGAGGTGATCTCTGCTCGAAAGAGAAATAGAAATATTTCTCCGGGCC 1860
Db 438 ATCAATAAGATATTTCTTCCCTT-----TACTCATAAAGTTGTCAGCTTTGGCC 489
QY 1861 AGGCGTGTGGCTATGCTGTAATCCAGCACTTTGGAGGCCCAAGCATGTGTGATCAC 1920
Db 490 GGGCGGATGGCTCACGCTGTTATCTAGTACTTTGGAGGCTCAGCGGCGCAGATCAC 549
QY 1921 CTGAGGTGAGGATTTCAAAACAGCTGGCCAACTGTTGAAACCCCTCTCTACTAAA 1980
Db 550 CTGAGGTGAGGATTTGTACAGCTGGCCAACTGTTGAAACCCCTCTCTACTAAA 609
QY 1981 ATACAAAAATTTAGTAAAGTGTGGCGCATGCTGTAATCCCACTTCTGGAGGCT 2040
Db 610 ATAC--AAACTAGCCAGTGTGTGTGGCCACCTGCGAGTCTTCCGAGGCT 667
QY 2041 GAGGAGGAGATTTCTTGAACCCGGAGGAGAGGTTGAGTGAAGCAGATCACCA 2100
Db 668 GAGGAGGAGATTTCAACCCGAGGAGGAGGTTGAGTGAAGCAGATTTGCGA 727
QY 2101 CTGCACTCCGCTGGGAGGAGGAGGAGTCTCTCTCAAAAAACAAAAACAAAGA 2160
Db 728 TCACACTCCGCTGGAGGAGGAGGAGGAGTCTCTCTCAAAAAACAAAAACAAAGA 786
QY 2161 ATTAAGCAAAATTTAGCATTTGACAGAGAACCTGAAGGGGGTTCAGACGATCAC 2220
Db 787 TGTAGCTTAAATAATCTCTCAGAGAGGATTTCTTTGCTCTAAACCTTCTCTTT 846
QY 2221 TGTGCCCATGCAAGTACTCTGAGGAGGATGCTG 2255
Db 847 TTTAAWCTGTGTGGCGGTTTGGAGGGAACAG 881
```

```
RESULT 14
BU616025
LOCUS
DEFINITION
  BU616025
  UI-H-DF0-ben-m-20-0-UI.s1 NCI CGAP DF0 Homo sapiens cDNA clone
  UI-H-DF0-ben-m-20-0-UI 3', mRNA sequence.
ACCESSION
  BU616025
VERSION
  BU616025.1
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 693)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
JOURNAL
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Dr. Jose Mercuende
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
```

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 9-305, >ALU (matched complement) 326-387, >L1MB7#LINE/L1
(matched complement) 388-676, >ALU
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

Location/Qualifiers

1..693

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-H-DF0-ben-m-20-0-UI"

/tissue_type="Subchondral Bone"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP DF0"

/note="Organ: Bone; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

NCI CGAP DF0 is a cDNA library containing the following

tissue(s): Subchondral Bone. The library was constructed

according to Bonaldo, Lennon and Soares, Genome Research,

6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT73-Pac

vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tag for this library is GTTAGCGTC.

TAG_TISSUE=Subchondral bone

TAG_LIB=UI-H-DF0

TAG_SEQ=GTTAGCGTC"

ORIGIN

```
Query Match      6.5%; Score 259.2; DB 13; Length 693;
Best Local Similarity 64.7%; Pred. No. 3.6e-30;
Matches 465; Conservative 0; Mismatches 216; Indels 38; Gaps 4;

QY 1441 TTTTCTTTTCTTTTGTGAGAAAGATCTTGTCTTTTCACTTGTGCTGGCTGAGGCGATGGTG 1500
Db 1 TTTTCTTTTCTTTTGTGAGAAAGATCTTGTCTTTTCACTTGTGCTGGCTGAGGCGATGGTG 60
QY 1501 TGATCTAGCTCAGTGCACATTTTCCCTCCCGGTTTCAAGCAATTTCTCTGCTTCAAGCT 1560
Db 61 CAATCACAGTTCACAGCAGCTGTAACTTGTGGCTCAAGCGGTCTCCACCTTAGCTT 120
QY 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCCACTGCCCCAGCTGATTTT---GTATTTT 1617
Db 121 CCCGGTAACTGGAACGTGAGTGTGCCACTATGCCAGCTAATTTTAAAAAAATTT 180
QY 1618 TAGTAGATGGGATTTTCACTTTTGTGCCAAGCTGTCTCAAACTTTTGTCTGCATAA 1677
Db 181 TTATAGATAGAGTCTCTCTATGTATACCTAAGCTGATCTCAAACTCTCTGG----- 231
QY 1678 TTGTTGTAATATTGTTCTTTTGTGAGGTAGGGCCCCCAGACCAAAAAATAAATCT 1737
Db 232 -----GCTCAGGCGAGTCTCTCTGNTCCGGCTCCCAAGTGTGGGATTAATAG 281
QY 1738 TAGAATCAAAATCAGTGTGTTGGTTTGACCTGTCACTTGAGAACCAAGTGTGACCG 1797
Db 282 GGGTGTAGGCCACACACACCCAGCCCAATACTGTTTTTAAAAAAGCTTAAT----- 333
QY 1798 GGCCTCAGGAGTAGAGGTGATCTCTGCTCGAAGAGAAATAGAAATATTTCTCCGG 1857
Db 334 -----TGAGACATACTTCAATACCATACCAATTTACATATTTAAAGTGTCAAAATCAG 386
QY 1858 GCCAGGCTGTGGCTCATGCTGTAATCCAGCACTTTGGAGGCCAAGGATCTGGAT 1917
Db 387 TCCGGGCGATGGCTCATGCTGTAATCCAGCACTTTGGAAGGCCGAGGAGTGGAT 446
QY 1918 CACCTGAGGTGAGGATTTCAAAACCAAGCTGGCCCAACATGGTGAACCCCTCTCTACTA 1977
```

```
Db 447 CACCTGAGCGCGGAGTTTCAGACCGGCTGGCCAAACATGGTGAACCCCGTCTCTACTA 506
QY 1978 AAAATACAAAAAATAGTCTAAGTGTGGTGGCGCATGCTGTAATCCACAGCTACTTGGGAG 2037
Db 507 AAAAT-CCAAAAATTAGTGGTGTGATGGCGTGACCTGTGATCCCAAGTTACTCAGGAG 565
QY 2038 GGTGAGGCGAGGAGAAATTTCTTGAACCCGCGGAGGCGAGAGTTGTCAGTGAAGCGAGATCACA 2097
Db 566 GCTGAGGCGAGGAGAAATCTCTGGACCTGGGAGCGGAGGTTGTCAGTGAAGCGAGATCACA 625
QY 2098 CCACGTGCACTCCAGCTCGGGGAGGAGGAGGAGCTTCTCTCAAAAAAACAACAA 2156
Db 626 CCACCTTACTCAAGCTCGGGTGACACAGCGAGATTTCTGTCTCANAAAAAACAACAA 684

RESULT 15
AF101960 Human Homo sapiens genomic clone pTWB59.14, genomic survey
LOCUS AF101960 2971 bp DNA linear GSS 06-NOV-2000
DEFINITION AF101960 Human Homo sapiens genomic clone pTWB59.14, genomic survey
sequence.
ACCESSION AF101960
VERSION AF101960.1 GI:4193786
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2971)
AUTHORS Bepler,G., O'Brian,K.C., Kim,Y.C., Schreiber,G. and Pitterle,D.M.
TITLE A 1.4-Mb high-resolution physical map and config. of chromosome
segment 11p15.5 and genes in the LOH11A metastasis suppressor
region
JOURNAL Genomics 55 (2), 164-175 (1999)
MEDLINE 99134294
PUBMED 9913563
COMMENT Contact: Bepler G
Medicine and Radiology
Duke University Medical Center
Box 2610, MSRB, Room 117, Durham, NC 27710, USA
part of a 1.4 megabase contig including the LOH11A metastasis
suppressor region Bin T
Class: unknown.
FEATURES             Location/Qualifiers
     source            1. .2971
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /map="11p15.5"
                        /clone="pTWB59.14"
                        /clone_lib="Human"

ORIGIN
Query Match          6.5%; Score 255.8; DB 28; Length 2971;
Best Local Similarity 65.4%; Pred. No. 5e-30;
Matches 478; Conservative 0; Mismatches 177; Indels 76; Gaps 4;
QY 1442 TTTTATTTTATTTTGGAGAAAGAGTCTGTCTTTTACCCCTGGCTGGAGGCGAGTGGTGT 1501
Db 478 TTTATTTATTTTCCGAGACAGTGTCTCACTGTGTTGCCAGACTGGAGTGCAGTGGCAC 537
QY 1502 GATCTCAGCTCAGTCACTTCGCTCCCGGGTTCAGCAATTCCTGCTTCAGCCTC 1561
Db 538 AATCTTGGCTCAGTCAACCTCCGCTCCCTGGGTTCAAGCGATTCTCCTGCCCTCAGCCTC 597
QY 1562 CCAAGTAGCTGGGACTACAGGCACTTCCACCATGCCAGCTGATTTTGTATTTTACT 1621
Db 598 CCAAGTAGCTGGGATTACAGGCGCCGCCACCAACAGCTAATTTTGTATTTTACT 657
QY 1622 AGAGATGGGATTTCACCTTTTGTGGCCAAAGCTGGTCTCAAACTTTTGTCTGCATATAATGT 1681
Db 658 AGAGATGAGGTTTTCACAAAGTTGGCCAGGCTGTCTCAAACTCTT----- 702
```

```
QY 1682 TGTAACATATTGTTCTTTTGTGAGGTAGGGCCCGCCAGACCAAAAAATAAATCTTAGA 1741
Db 703 -----GACCTCGTGATCCACCCACTCAGCCTCCCA 733
QY 1742 ATCCAAATACAGTGTGTTGGTTTGACCACTGTCTTGAGAACCAACAGTGTGACCGAGGCC 1801
Db 734 AAGTGCTGGGATTACAGGCATAAGCCACCGCATCTGGCCCCAACATTATAT----- 783
QY 1802 TCAGGAGTAGAGGTGATCTCTGCTCGAAAGAGAAATAGAAATATTTCTCCGGGCA 1861
Db 784 -----TTTATGGTGAAGACTGAATGTTTCTTAAGATCAGAAACAGCCA 827
QY 1862 GGCCTGGTGGTCAATGCTGTAATCCACAGCACTTTGGAGGCCAAGGCATGTGATCACC 1921
Db 828 GGCACAGTGGTCAAGCTGTGTAATCTGGCACTTTGGAGGCCGAGCGGCAGATCAC- 886
QY 1922 TGAGGTGAGGAGTTCAAAACAGGCTGGCCAAACATGTTGAACCCCGGAGGTCAGTGAAGCGAGATCACA 2097
Db 887 -GAGGTGAGGAGATCGAGACCATCTCTGGCTAACACAGGTGAACCCCGTCTCTACTAAAAA 945
QY 1982 T----ACAAAAAATTAGCTAAGTGTGGTGGCGCATGCTGTAATCCCGAGCTACTTGGGAG 2037
Db 946 TAAAAATAAAAAATTAGCTGGTGCAGTGGTGGCACCTGTAGTCCAGCTACTCGGGAG 1005
QY 2038 GGTGAGCGAGGAGAAATTTCTTGAACCCCGGAGGCGAGAGTTGTCAGTGAAGCGAGATCACA 2097
Db 1006 GCTGAGCGCGGAGAAATGGTGTGAACCCAGAGGCGGAGCTTGCAGTGAAGCGAGATCATG 1065
QY 2098 CCACTGCACCTCCAGCTTGGGGGAGAGGAGAGCTTCTCTCAAAAAACAACAAAAA 2157
Db 1066 CCACTGCACCTCCAGCTTGGGCGACAGGGCGAGACTCCGACTCAAAAAAAGAGATTA 1125
QY 2158 AGAATTAAAGCA 2168
Db 1126 GAAATGAACA 1136
```

Search completed: May 13, 2004, 17:08:27
Job time : 14558.7 secs